Error analysis for a class of numerical differentiator

Da-Yan Liu, Dayan Liu, Olivier Gibaru, Wilfrid Perruquetti

Error analysis for a class of numerical differentiator

Dayan Liu, Olivier Gibaru, Wilfrid Perruquetti[1]

Abstract This report is devoted to derivatives estimations. Contrary to the Tikhonov's regularization procedure, we use a recent algebraic framework which involves finally a projection into the Jacobi polynomial basis so as to estimate these derivatives from noisy data. No information about the statistical properties of the noise is required. We give some results concerning the choice of the parameters of this method so as to minimize the noise error contribution and the approximation errors. Moreover, two new central estimators based on such algebraic differentiation techniques are introduced. A comparison is done between these estimations and some of the improved classical numerical differentiation schemes.

Keywords Numerical differentiation, Operational calculus, Jacobi orthogonal polynomials, Numerical integration, Gaussian noise, Error bound

Mathematics Subject Classification (2000) 65D25, 44A40, 44A10, 33C45, 65D30, 60H40, 41A58

1 Introduction

Since the seminal paper by Diop & Fliess ([16] see also [4]), observation theory and identifiability are closely linked to numerical differentiation scheme. Indeed, a non-linear input-output system is observable if, and only if, any state variable is a differential function of the control and output variables, *i.e.* a function of those variables and their derivatives up to some finite order.

Recent algebraic parametric estimation technics for linear systems [21, 20] have been extended to various problems in signal processing (see, e.g., [17, 33, 36, 37, 44, 45, 46]). Let us emphasize that those methods, which are algebraic and non-asymptotic, exhibit good robustness properties with respect to corrupting noises, without the need of knowing their statistical properties (See [14, 15] for more theoretical details). The robustness properties have already been confirmed by numerous computer simulations and several laboratory experiments. It appears that these technics can also be used to derive numerical differentiation algorithms exhibiting similar properties (see [34, 35]). Such technics are used in [18, 19, 4] for state estimation.

Numerical differentiation is concerned with the estimation of derivatives of noisy time signals. This problem has attracted a lot of attention from different points of view

- observer design in the control literature (see [8, 10, 24, 25, 31, 42]),
- digital filter in signal processing (see [1, 7, 11, 38, 41]),

for on-line applications which are alternative solutions to the very classical one, based on least-squares polynomial fitting or (spline) interpolation mostly used in off-line applications ([13, 26]).

In recent papers [34, 35], numerical differentiation is revised using an algebraic framework of parameter estimation. To start with, let $y(t) = x(t) + \varpi(t)$ be a noisy observation on a finite time interval I (an open interval of $\mathbb{R}^+$) of a real valued smooth signal x, the successive derivatives of which we want to estimate, and $\varpi(t)$ denotes the noise. Considering x as an analytic function on I, the Taylor series expansion of x at t_0 is:

$$x(t_0 \pm t) = \sum_{i=0}^{+\infty} \frac{(\pm t)^i}{i!} x^{(i)}(t_0),\ \forall t_0 \in I \text{ and } t \geq 0 \text{ as soon as } t + t_0 \in I \ \text{ (resp. } t_0 - t \in I). \tag{1}$$

[1]D. Liu, O. Gibaru, W. Perruquetti are with Équipe Projet ALIEN, INRIA Lille-Nord Europe, Parc Scientifique de la Haute Borne 40, avenue Halley Bât.A, Park Plaza, 59650 Villeneuve d'Ascq, France. O. Gibaru is with Arts et Métiers ParisTech centre de Lille, L2MA: Mathématiques Appliquées et mécanique de précision, 8 Boulevard Louis XIV, 59046 Lille Cedex, France, W. Perruquetti is with LAGIS (CNRS, UMR 8146), École Centrale de Lille, BP 48, Cité Scientifique, 59650 Villeneuve d'Ascq, France and `dayan.liu@inria.fr`, `olivier.gibaru@ensam.eu`, `wilfrid.perruquetti@inria.fr`

Contrary to the numerical differentiation method in [23], where the Taylor series expansions of x are used and an elimination technic is applied without any error analysis due to the bias term or the noise (see Remark 1 for more details), we achieve the estimations of the derivatives of x by using an algebraic framework. This approach allows us to study in detail the bias term error and the noise error contribution.

Using a representation in the operational calculus framework of the $N^{th} \geq n$ (n is the order de the derivative of x which we want to estimate) order truncation of the Taylor series expansion of (1) and an annihilator (which is a linear differential operator) of a specific form

$$\Pi_{k,\mu}^{N,n} = \frac{1}{s^{N+1+\mu}} \cdot \frac{d^{n+k}}{ds^{n+k}} \cdot \frac{1}{s} \cdot \frac{d^{N-n}}{ds^{N-n}} \cdot s^{N+1},$$

the authors Mboub, Join and Fliess (see [34, 35]) obtained non asymptotic estimations of the n^{th} order derivative of $x(t)$. These expressions are based on iterated integrals of a polynomial multiplied by the noisy observation signal. They depend on four parameters k, μ, N, T as follows:

$$\tilde{x}_{t_{0\pm}}^{(n)}(k, \mu, \pm T, N) = \int_0^1 p_{k,\mu,\pm T,n,N}(\tau)\, y(\pm T\tau + t_0) d\tau, \tag{2}$$

where $\tilde{x}_{t_{0-}}^{(n)}(k, \mu, -T, N)$ is the causal estimator, $\tilde{x}_{t_{0+}}^{(n)}(k, \mu, T, N)$ is the anti-causal estimator respectively and $p_{k,\mu,\pm T,n,N}$ are polynomials. If the noisy function y is integrable in (2), then by applying a change of variables in (2): $\pm T\tau + t_0 \to \tau$, we obtain a Fredholm integral equation of the first kind where the variables of the kernel are τ and t_0. A similar formula is shown in [9] where the numerical differentiation method is equivalent to solving a particular Fredholm integral equation of the first kind by applying the Tikhonov regularization method, and the solution considered as a derivative estimator is also an integral estimator. The hypothesis on the noise is more general in the estimations defined in (2) than the one used in [9] and in [32] where the noise is assumed to be bounded. However, in [32] where the Tikhonov regularization method is also employed, the author studies the case for unstructured and noisy data in a d-dimension space, which is out of the scope of this paper.

Moreover, it was shown that the estimations defined in (2) can be viewed as an affine combination of minimal estimations obtained for the Taylor series truncation at order n (with $n < N$):

$$\tilde{x}_{t_{0\pm}}^{(n)}(k, \mu, \pm T, N) = \sum_{l=0}^{N-n} \lambda_l \tilde{x}_{t_{0\pm}}^{(n)}(k_l, \mu_l, \pm T, n), \quad \lambda_l \in \mathbb{Q}. \tag{3}$$

It was shown that if $\lambda_l \in \mathbb{R}$, then $\tilde{x}_{t_{0\pm}}^{(n)}(k, \mu, \pm T, N)$ give us general formulas of causal estimator and anti-causal estimator, which correspond to the projections of the noisy signal into the Jacobi polynomial basis. For these estimators, there are two sources of error: the bias term errors which comme from the truncation of the Taylor series expansion, and the noise error contributions. The bias terms error can be significantly reduced by allowing the estimated signal to be shifted by a small delay $T\xi$, where ξ is the smallest root of a given Jacobi polynomial. However, if one wants to minimize the noise error contributions, then one has to select the λ_l being equal to the inverse of the difference $N - n + 1$. Thus, the set of parameters has to be selected to achieve a compromise between reducing the bias term and reducing the noise error contribution. Nevertheless, a weakness of the above method was a lack of any error analysis, while they were being implemented in practice. Thus one aim of this paper is to provide a guide to parameters selection in order to achieve this compromise.

The paper is organized as follows:

Section 2 gives a general smoothness assumption on the signal x of which successive derivative has to be estimated. Using a truncation of the Taylor expansion of an affine combination of x (to be replaced after by the noisy measured signal), a time dependent signal is obtained to be used for the computation of the desired estimation. Then, this signal is passed to the operational domain by using Laplace Transform. A general differential operator parameterized by a set is applied with an elimination technic in order to obtain, in the time domain, an integral estimator of the n^{th} order derivative of x. Sufficient and necessary conditions on the set of parameters involved in these operators are given to obtain such an annihilator and it is proven that such set of parameters is non empty. Let us stress that such technics are rooted in [34, 35], thus Section 3 firstly recalls the minimal estimators and the affine estimators introduced by M. MBoup, C. Join and M. Fliess in [34, 35]. For such estimators lower bounds and upper bounds for the bias term error are given. In order to reduce the bias term error, two new families of central estimators are given. Section 4 is devoted to the error analysis. This error is due to the noise, the numerical approximations of the integrals and the order of truncation of the Taylor expansion. The noise is assumed to be a sequence of independent random variables with the same expected value and the same variance. Thus the noise error contribution for the proposed estimators can be given thanks to the numerical integration method. Moreover, with the knowledge of the expected value and the variance of the noise and according to the Bienaymé-Chebyshev inequality, a lower bound and an upper bound for discrete noise error contribution are obtained. In particular, the exact expressions of the bounds for the noise error contributions are given for the first order derivative estimators in the case of a gaussian white noise. Section 5 compares the obtained estimators with some of the existing numerical differentiation methods, such as the averaged finite difference schemes and Savitzky-Golay differentiation scheme. For this, two criterions are used: the integral of the error square and the classical SNR.

2 Preliminaries

Let us start with $y(t) = x(t) + \varpi(t)$ a noisy observation on a finite time open interval $I \subset \mathbb{R}^+$ of a real valued smooth signal x, the successive derivatives of which we want to estimate, and $\varpi(t)$ denotes a noise. Let $n \in \mathbb{N}$, we are going to estimate the n^{th} order derivative of x.

2.1 Operational domain

Let us ignore the noise $\varpi(t)$ for the moment. Assume that $x(t)$ is an analytic function on I. For $t_0 \in I$, let us introduce the following function

$$X(t) = \sum_{i=0}^{L} a_i \, x(t_0 + \beta_i t), \tag{4}$$

where $L \in \mathbb{N}, a_i \in \mathbb{R}^*, \beta_i \in \mathbb{R}^*$, $\beta_0 < \beta_1 < \cdots < \beta_L$ and $t \in D := \{t \in \mathbb{R}^+; \forall i \in \{1, ..., L\}, t_0 + \beta_i t \in I\}$. This analytic function $X(t)$ will be used to perform any derivatives estimations of x at point t_0 in only one general framework. Actually, if all the $\beta_i < 0$ (resp. $\beta_i > 0$), then we will obtain causal estimators (resp. anti-causal estimators). In the other cases, we will obtain "finite difference" type estimators. Consequently, $X(t)$ is also an analytic function on D. The Taylor series expansion of X at t_0 is given by

$$\forall t \in D,\ X(t) = \sum_{i=0}^{L} a_i \sum_{j=0}^{+\infty} \frac{(\beta_i t)^j}{j!} x^{(j)}(t_0). \tag{5}$$

For $N \in \mathbb{N} \geq n$, we consider the following truncated Taylor series expansion of X on $\mathbb{R}^+$:

$$\forall t \in \mathbb{R}^+,\ X_N(t) = \sum_{i=0}^{L} a_i \sum_{j=0}^{N} \frac{(\beta_i t)^j}{j!} x^{(j)}(t_0) = \sum_{j=0}^{N} \left(\sum_{i=0}^{L} a_i \beta_i^j \right) \frac{t^j}{j!} x^{(j)}(t_0). \tag{6}$$

Since X_N is a polynomial defined on $\mathbb{R}^+$ of degree N, we can apply the Laplace transform to (6) (s being the Laplace variable):

$$\hat{X}_N(s) = \sum_{j=0}^{N} c_j s^{-(j+1)} x^{(j)}(t_0), \tag{7}$$

where $\hat{X}_N(s)$ is the Laplace transform of $X_N(t)$, $c_j = \sum_{i=0}^{L} a_i \beta_i^j$, and c_n is supposed to be different from zero.

In all the sequel, the Laplace transform of a signal $u(t)$ will be denoted by $\hat{u}(s)$. To simplify the notation, the argument s will be dropped and we write it as $\hat{u}$ for short.

2.2 Annihilators

The basic step towards the estimation of $x^{(n)}(t)$, for $t \geq 0$, is the estimation of the coefficient $x^{(n)}(t_0)$ from the observation $y(t)$. All the terms $c_j s^{-(j+1)} x^{(j)}(t_0)$ in (7) with $j \neq n$, are consequently considered as undesired terms which we proceed to annihilate. For this, it suffices to find a linear differential operator of the form

$$\Pi = \sum_{\text{finite}} \left(\prod_{\text{finite}} \varrho_l(s) \frac{d^l}{ds^l} \right), \quad \varrho_l(s) \in \mathbb{C}(s), \tag{8}$$

such that

$$\Pi \left(\hat{X}_N(s) \right) = \varrho(s)\, x^{(n)}(t_0), \tag{9}$$

for some rational function $\varrho(s) \in \mathbb{C}(s)$. Such a linear differential operator is subsequently called an *annihilator* for $x^{(n)}(t_0)$ (see [35]).

When the sum in (8) is reduced to a single term, we obtain a particular case of such linear differential operator which is a finite product of length $\Theta \in \mathbb{N}$. If for all indexes l, the rational function $\varrho_l(s)$ is of the following form $\varrho_l(s) = \frac{1}{s^{m_l}}$, then the linear differential operator defined by (8) can be parameterized by the set $E = \{(n_l, m_l)\}_{l=1}^{\Theta}$:

$$\Pi_E = \prod_{l=1}^{\Theta} \frac{1}{s^{m_l}} \frac{d^{n_l}}{ds^{n_l}} = \frac{1}{s^{m_1}} \frac{d^{n_1}}{ds^{n_1}} \cdots \frac{1}{s^{m_\Theta}} \frac{d^{n_\Theta}}{ds^{n_\Theta}}. \tag{10}$$

Note that $m_l \in \mathbb{Z}^*$ for $l = 1, \ldots, \Theta$, and $n_l \in \mathbb{N}^*$ for $l = 1, \ldots, \Theta - 1$, except for $n_\Theta \in \mathbb{N}$. For such an operator, in the following proposition, we are going to obtain explicitly the value of $\Pi_E \left(\hat{X}_N(s) \right)$, where $\hat{X}_N(s)$ is defined by (7). We also give conditions on the integers m_l and n_l, such that Π_E preserves only the $x^{(n)}(t_0)$ term such as the case in (9).

Proposition 2.1 *Let $\hat{X}_N(s)$ be defined by (7) and Π_E be the linear differential operator defined by (10). If E satisfies the following conditions:*

(C1): $\forall l \in \{1, \ldots, \Theta - 1\}$: either $n + 1 + r_l > 0$ or $n + 1 + r_l \leq -n_l$ is true,

(C2): for each $j \in \mathbf{J} = \{k; k \in \{0, \ldots, n-1, n+1, \ldots, N\}, c_k x^{(k)}(t_0) \neq 0\}$, there exists a $l_j \in \{1, ..., \Theta - 1\}$, such that $0 \leq -(j+1) - r_{l_j} < n_{l_j}$,

with $r_l = \sum_{i=l+1}^{\Theta} n_i + m_i$ for $l = 0, \ldots, \Theta - 1$ and $r_\Theta = 0$. Then Π_E is an annihilator and

$$\Pi_E \left(\hat{X}_N(s) \right) = c_n x^{(n)}(t_0) \frac{\hat{c}}{s^{n+1+r_0}}, \tag{11}$$

where $\hat{c} = \prod_{l=1}^{\Theta} \hat{c}_l$, with

$$\hat{c}_l = \begin{cases} \frac{(-1)^{n_l} (n_l + n + r_l)!}{(n + r_l)!}, & \text{if } n + 1 + r_l > 0, \\ \frac{(|n+1+r_l|)!}{(|n+1+r_l| - n_l)!}, & \text{if } n + 1 + r_l \leq -n_l. \end{cases} \tag{12}$$

Moreover, if $c_n x^{(n)}(t_0) \neq 0$, then the conditions $(C1)$ and $(C2)$ are also necessary.

Remark 1 *The condition $(C2)$ is used to annihilate all the undesired terms: $c_j s^{-(j+1)} x^{(j)}(t_0)$ in (7) for all $j \in \mathbf{J}$, and the condition $(C1)$ is used to keep the term $c_n s^{-(n+1)} x^{(n)}(t_0)$. A similar technic was used in [23] in order to estimate the n^{th} order derivative of a smooth function x. The Taylor series expansions at a point t_0 in different intervals for x were given. Then an affine combination of these series could annihilate all the undesired terms which contain $x^{(i)}(t_0)$ for $i < n$. All the remainder terms which contain $x^{(i)}(t_0)$ for $i > n$ were considered as a bias term error. However, no error analysis was conducted in [23].*

Proof: By applying the linear differential operator as defined by (10) to the right hand side of (7), one obtains

$$\Pi_E\left(\sum_{j=0}^{N} c_j s^{-(j+1)} x^{(j)}(t_0)\right) = \sum_{j=0}^{N} c_j x^{(j)}(t_0)\ \Pi_E(s^{-(j+1)}).$$

• Sufficiency: The computation is divided into two parts: one concerning the term $\Pi_E(s^{-(j+1)})$ with $j = n$, and one concerning the others with $j \neq n$. We will show that the annihilator kills these terms with $j \neq n$. Recall firstly the following formula: for $k \in \mathbb{N}$ and $m \in \mathbb{Z}^*$, $\frac{d^k(s^m)}{ds^k}$ is given by

$$\frac{m!}{(m-k)!} s^{m-k} \quad \text{if } 0 \leq k < m, \tag{13a}$$

$$0 \quad \text{if } 0 \leq m < k, \tag{13b}$$

$$\frac{(-1)^k (k-m-1)!}{(-m-1)!} s^{m-k} \quad \text{if } m < 0 \leq k. \tag{13c}$$

Computation of $\Pi_E(s^{-(n+1)})$: By induction, we want to prove that:

$$\prod_{l=J}^{\Theta} \frac{1}{s^{m_l}} \frac{d^{n_l}}{ds^{n_l}} s^{-(n+1)} = \frac{1}{s^{n+1+r_{J-1}}} \prod_{l=J}^{\Theta} \hat{c}_l \tag{14}$$

holds for any $J \in \{1, \ldots, \Theta\}$, where $\hat{c}_l = \frac{(-1)^{n_l}(n_l+n+r_l)!}{(n+r_l)!}$, if $n+1+r_l > 0$, and $\hat{c}_l = \frac{(|n+1+r_l|)!}{(|n+1+r_l|-n_l)!}$, if $n+1+r_l \leq -n_l$.
Initial step: when $J = \Theta$, using (13c) one obtains

$$\frac{1}{s^{m_\Theta}} \frac{d^{n_\Theta}}{ds^{n_\Theta}} s^{-(n+1)} = \frac{\hat{c}_\Theta}{s^{n+1+r_{\Theta-1}}}, \quad \text{with } \hat{c}_\Theta = \frac{(-1)^{n_\Theta}(n_\Theta + n)!}{n!}. \tag{15}$$

Assume now that (14) holds for $1 < J \leq \Theta$, this leads to

$$\prod_{l=J-1}^{\Theta} \frac{1}{s^{m_l}} \frac{d^{n_l}}{ds^{n_l}} s^{-(n+1)} = \frac{1}{s^{m_{J-1}}} \frac{d^{n_{J-1}}}{ds^{n_{J-1}}} \cdot \frac{1}{s^{n+1+r_{J-1}}} \prod_{l=J}^{\Theta} \hat{c}_l.$$

We distinguish now the two following cases in condition $(C1)$:

1. If $n + 1 + r_{J-1} > 0$, then using (13c) yields

$$\prod_{l=J-1}^{\Theta} \frac{1}{s^{m_l}} \frac{d^{n_l}}{ds^{n_l}} s^{-(n+1)} = \frac{1}{s^{n+1+r_{J-2}}} \prod_{l=J-1}^{\Theta} \hat{c}_l, \text{ with } \hat{c}_{J-1} = \frac{(-1)^{n_{J-1}}(n_{J-1} + n + r_{J-1})!}{(n + r_{J-1})!}.$$

2. If $n + 1 + r_{J-1} \leq -n_{J-1}$, then using (13a) yields

$$\prod_{l=J-1}^{\Theta} \frac{1}{s^{m_l}} \frac{d^{n_l}}{ds^{n_l}} s^{-(n+1)} = \frac{1}{s^{n+1+r_{J-2}}} \prod_{l=J-1}^{\Theta} \hat{c}_l, \text{ with } \hat{c}_{J-1} = \frac{(|n + 1 + r_{J-1}|)!}{(|n + 1 + r_{J-1}| - n_{J-1})!}.$$

Consequently, we can conclude that (14) is true for $J - 1$. Thus by induction, (14) is true for any $J \in \{1, \ldots, \Theta\}$.
Computation of $\Pi_E(s^{-(j+1)}), j \in \mathbf{J}$:
For any $l_j \in \{1, ..., \Theta - 1\}$, we have

$$\prod_{l=l_j+1}^{\Theta} \frac{1}{s^{m_l}} \frac{d^{n_l}}{ds^{n_l}} s^{-(j+1)} = \frac{\tilde{c}_{l_j}}{s^{j+1+r_{l_j}}}, \text{ with } \tilde{c}_{l_j} \in \mathbb{Q}.$$

From condition $(C2)$, we know that $0 \leq -(j+1) - r_{l_j} < n_{l_j}$, then we obtain

$$\Pi_E\left(s^{-(j+1)}\right) = \frac{d^{n_{l_j}}}{ds^{n_{l_j}}}\left(\prod_{l=l_j+1}^{\Theta}\frac{1}{s^{m_l}}\frac{d^{n_l}}{ds^{n_l}}s^{-(j+1)}\right) = 0. \tag{16}$$

Finally, we conclude that

$$\Pi_E\left(\sum_{j=0}^{N} c_j s^{-(j+1)} x^{(j)}(t_0)\right) = c_n x^{(n)}(t_0)\Pi_E(s^{-(n+1)}) = c_n x^{(n)}(t_0)\frac{\hat{c}}{s^{n+1+r_0}}.$$

Thus, we have shown that conditions $(C1)$ and $(C2)$ are sufficient conditions for (11) to hold.
• Necessity: We are going to prove that the conditions $(C1)$ and $(C2)$ are also necessary as soon as $c_n x^{(n)}(t_0) \neq 0$. To do so, let us assume that (11) is true, then

$$\sum_{j=0}^{N} c_j x^{(j)}(t_0)\Pi_E\left(s^{-(j+1)}\right) = c_n x^{(n)}(t_0)\frac{\hat{c}}{s^{n+1+r_0}}. \tag{17}$$

Using similar calculations leading to (14), we obtain (without using conditions $(C1)$ and $(C2)$): for $0 \leq j \leq N$, $\Pi_E\left(s^{-(j+1)}\right) = \frac{\bar{c}_j}{s^{j+1+r_0}}$, with $\bar{c}_j \in \mathbb{Q}$. Thus, (17) becomes

$$\forall s \in \mathbb{C} \text{ with } \Re(s) > 0, \quad \sum_{j\in\mathbf{J}} c_j x^{(j)}(t_0)\frac{\bar{c}_j}{s^{j+1+r_0}} + c_n x^{(n)}(t_0)\frac{\hat{c}-\bar{c}_n}{s^{n+1+r_0}} = 0. \tag{18}$$

Therefore, it can be seen that $c_n x^{(n)}(t_0)(\hat{c}-\bar{c}_n) = 0$ and $c_j x^{(j)}(t_0)\bar{c}_j = 0$ for all $j \in \mathbf{J}$. As $c_j x^{(j)}(t_0) \neq 0$ for all $j \in \mathbf{J} \cup \{n\}$, we have $\hat{c} = \bar{c}_n$ and $\bar{c}_j = 0$ for all $j \in \mathbf{J}$. Consequently, $\Pi_E\left(s^{-(n+1)}\right) = \frac{\hat{c}}{s^{n+1+r_0}}$ and $\Pi_E\left(s^{-(j+1)}\right) = 0$ for all $j \in \mathbf{J}$. Since $\Pi_E\left(s^{-(n+1)}\right) \neq 0$, $s^{-(n+1)}$ is not annihilated after each derivation, we can conclude that $\forall l \in \{1,\ldots,\Theta-1\}$, either $n+1+r_l > 0$ or $n+1+r_l \leq -n_l$ is true. On the other hand, $s^{-(j+1)}$ is annihilated for all $j \in \mathbf{J}$. Hence, there exists a $l_j \in \{1,...,\Theta-1\}$, such that $0 < -(j+1) - r_{l_j} < n_{l_j}$. ■

In the following proposition, we will give some new conditions on the set E such that in the time domain the estimator of $x^{(n)}(t_0)$ only depends on a unique integral of the measured signal. Before doing so, we propose the following lemma.

Lemma 2.2 *Let $\hat{f}$ be the Laplace transform of an analytic function f defined by $\mathbb{R}^+$ and Π_E be an operator defined by (10). We split each m_l into two terms: $m_l = \hat{m}_l + \bar{m}_l$ with $\hat{m}_l \in \mathbb{Z}^*$ and $\bar{m}_l \in \mathbb{Z}$ for $l = 1,\cdots,\Theta$. Let $\mathbf{j} = (j_1,...,j_\Theta)$ be a multi-index of length Θ and $\bar{E}_{\mathbf{j}} = \{(n_l - j_l, \bar{m}_l)\}_{l=1}^{\Theta}$ be a subset of $\mathbb{N}\times\mathbb{Z}$, then Π_E can be written as follow*

$$\Pi_E\left(\hat{f}(s)\right) = \sum_{j_\Theta=0}^{I_\Theta}\cdots\sum_{j_1=0}^{I_1} C_1\frac{1}{s^{\gamma_1}}\Pi_{\bar{E}_{\mathbf{j}}}\left(\hat{f}(s)\right), \tag{19}$$

where $\gamma_l = \sum_{i=l}^{\Theta}\hat{m}_i + j_i$ and $C_l = \prod_{i=l}^{\Theta} e_{j_i}$ for $l = 1,\cdots,\Theta$. The values

$$e_{j_l} = \begin{cases} \binom{n_l}{j_l}\frac{(-1)^{j_l}(j_l+\gamma_{l+1}-1)!}{(\gamma_{l+1}-1)!}, & \text{if } \gamma_{l+1} > 0, \\ \binom{n_l}{j_l}\frac{|\gamma_{l+1}|!}{(|\gamma_{l+1}|-j_l)!}, & \text{else}, \end{cases} \quad \text{and} \quad I_l = \begin{cases} n_l, & \text{if } \gamma_{l+1} > 0, \\ \min(n_l, |\gamma_{l+1}|), & \text{else}, \end{cases}$$

for $l = 1,\cdots,\Theta-1$. For $l = \Theta$ we have $I_\Theta = 0$, $e_{j_\Theta} = 1$.

Remark 2 *Here $\Pi_{\bar{E}_j}$ is an operator that can or not be also an annihilator. The criterion is that (9) holds for some $\rho(s)$.*

Proof: We are going to prove the following relation by induction: for $J = 1, \cdots, \Theta$,

$$\prod_{l=J}^{\Theta} \frac{1}{s^{m_l}} \frac{d^{n_l}}{ds^{n_l}} \left(\hat{f}(s)\right) = \sum_{j_\Theta=0}^{I_\Theta} \cdots \sum_{j_J=0}^{I_J} C_J \frac{1}{s^{\gamma_J}} \prod_{l=J}^{\Theta} \frac{1}{s^{\bar{m}_l}} \frac{d^{n_l-j_l}}{ds^{n_l-j_l}} \left(\hat{f}(s)\right). \tag{20}$$

Initial step: For $J = \Theta$, we have

$$\frac{1}{s^{m_\Theta}} \frac{d^{n_\Theta}}{ds^{n_\Theta}} \left(\hat{f}(s)\right) = \frac{1}{s^{\hat{m}_\Theta}} \frac{1}{s^{\bar{m}_\Theta}} \frac{d^{n_\Theta}}{ds^{n_\Theta}} \left(\hat{f}(s)\right).$$

Hence, the relation (20) is true for $J = \Theta$ with $I_\Theta = 0$, $C_\Theta = e_{j_\Theta} = 1$ and $\gamma_\Theta = \hat{m}_\Theta$.
Now assume that the relation (20) is true for $1 < J \le \Theta$, this leads to

$$\begin{aligned}
\prod_{l=J-1}^{\Theta} \frac{1}{s^{m_l}} \frac{d^{n_l}}{ds^{n_l}} \left(\hat{f}(s)\right) &= \frac{1}{s^{m_{J-1}}} \frac{d^{n_{J-1}}}{ds^{n_{J-1}}} \cdot \left(\sum_{j_\Theta=0}^{I_\Theta} \cdots \sum_{j_J=0}^{I_J} C_J \frac{1}{s^{\gamma_J}} \prod_{l=J}^{\Theta} \frac{1}{s^{\bar{m}_l}} \frac{d^{n_l-j_l}}{ds^{n_l-j_l}} \left(\hat{f}(s)\right) \right) \\
&= \sum_{j_\Theta=0}^{I_\Theta} \cdots \sum_{j_J=0}^{I_J} C_J \frac{1}{s^{m_{J-1}}} \frac{d^{n_{J-1}}}{ds^{n_{J-1}}} \cdot \left(\frac{1}{s^{\gamma_J}} \prod_{l=J}^{\Theta} \frac{1}{s^{\bar{m}_l}} \frac{d^{n_l-j_l}}{ds^{n_l-j_l}} \left(\hat{f}(s)\right) \right) \\
&= \sum_{j_\Theta=0}^{I_\Theta} \cdots \sum_{j_J=0}^{I_J} C_J \frac{1}{s^{m_{J-1}}} \sum_{j_{J-1}=0}^{n_{J-1}} \binom{n_{J-1}}{j_{J-1}} \frac{d^{j_{J-1}}}{ds^{j_{J-1}}} \left(\frac{1}{s^{\gamma_J}}\right) \frac{d^{n_{J-1}-j_{J-1}}}{ds^{n_{J-1}-j_{J-1}}} \left(\prod_{l=J}^{\Theta} \frac{1}{s^{\bar{m}_l}} \left(\hat{f}(s)\right) \right).
\end{aligned}$$

If $\gamma_J > 0$, we obtain

$$\begin{aligned}
&\prod_{l=J-1}^{\Theta} \frac{1}{s^{m_l}} \frac{d^{n_l}}{ds^{n_l}} \left(\hat{f}(s)\right) \\
&= \sum_{j_\Theta=0}^{I_\Theta} \cdots \sum_{j_J=0}^{I_J} C_J \sum_{j_{J-1}=0}^{n_{J-1}} \binom{n_{J-1}}{j_{J-1}} \frac{(-1)^{j_{J-1}}(j_{J-1}+\gamma_J-1)!}{(\gamma_J-1)!} \frac{1}{s^{m_{J-1}+\gamma_J+j_{J-1}}} \frac{d^{n_{J-1}-j_{J-1}}}{ds^{n_{J-1}-j_{J-1}}} \left(\prod_{l=J}^{\Theta} \frac{1}{s^{\bar{m}_l}} \frac{d^{n_l-j_l}}{ds^{n_l-j_l}} \left(\hat{f}(s)\right) \right) \\
&= \sum_{j_\Theta=0}^{I_\Theta} \cdots \sum_{j_{J-1}=0}^{I_{J-1}} C_{J-1} \frac{1}{s^{\gamma_{J-1}}} \prod_{l=J-1}^{\Theta} \frac{1}{s^{\bar{m}_l}} \frac{d^{n_l-j_l}}{ds^{n_l-j_l}} \left(\hat{f}(s)\right),
\end{aligned}$$

where $I_{J-1} = n_{J-1}$, $C_{J-1} = C_J \cdot e_{j_{J-1}}$, $e_{j_{J-1}} = \binom{n_{J-1}}{j_{J-1}} \frac{(-1)^{j_{J-1}}(j_{J-1}+\gamma_J-1)!}{(\gamma_J-1)!}$ and $\gamma_{J-1} = \gamma_J + \hat{m}_{J-1} + j_{J-1} = \sum_{i=J-1}^{\Theta} \hat{m}_i + j_i$.

If $\gamma_J \le 0$, we obtain

$$\begin{aligned}
&\prod_{l=J-1}^{\Theta} \frac{1}{s^{m_l}} \frac{d^{n_l}}{ds^{n_l}} \left(\hat{f}(s)\right) \\
&= \sum_{j_\Theta=0}^{I_\Theta} \cdots \sum_{j_J=0}^{I_J} C_J \sum_{j_{J-1}=0}^{\min(n_{J-1},|\gamma_J|)} \binom{n_{J-1}}{j_{J-1}} \frac{|\gamma_J|!}{(|\gamma_J|-j_{J-1})!} \frac{1}{s^{m_{J-1}+\gamma_J+j_{J-1}}} \frac{d^{n_{J-1}-j_{J-1}}}{ds^{n_{J-1}-j_{J-1}}} \left(\prod_{l=J}^{\Theta} \frac{1}{s^{\bar{m}_l}} \frac{d^{n_l-j_l}}{ds^{n_l-j_l}} \left(\hat{f}(s)\right) \right) \\
&= \sum_{j_\Theta=0}^{I_\Theta} \cdots \sum_{j_{J-1}=0}^{I_{J-1}} C_{J-1} \frac{1}{s^{\gamma_{J-1}}} \prod_{l=J-1}^{\Theta} \frac{1}{s^{\bar{m}_l}} \frac{d^{n_l-j_l}}{ds^{n_l-j_l}} \left(\hat{f}(s)\right),
\end{aligned}$$

where $I_{J-1} = \min(n_{J-1}, |\gamma_J|)$, $C_{J-1} = C_J \cdot e_{j_{J-1}}$, $e_{j_{J-1}} = \binom{n_{J-1}}{j_{J-1}} \frac{\gamma_J!}{(\gamma_J - j_{J-1})!}$ and $\gamma_{J-1} = \gamma_J + \hat{m}_{J-1} + j_{J-1} = \sum_{i=J-1}^{\Theta} \hat{m}_i + j_i$.

Consequently, we can conclude that the relation (20) is true for $J-1$. Thus by induction, (20) is true for any $J \in \{1, \ldots, \Theta\}$. Hence, (19) holds with $J = 1$. ∎

Proposition 2.3 *Let $\hat{X}_N(s)$ be defined by (7) and the annihilator Π_E be defined by (10), where the m_l satisfies the following condition ($C3$): $\sum_{l=1}^{\Theta} m_l \geq 1$. Then it yields*

$$\mathcal{L}^{-1}\left\{\Pi_E\left(\hat{X}_N(s)\right)\right\}(t) = \int_0^t p_{t,\Theta}(\tau) X_N(\tau) d\tau, \tag{21}$$

where $p_{t,\Theta}(\tau) = \sum_{j_\Theta=0}^{I_\Theta} \cdots \sum_{j_1=0}^{I_1} C_1 \frac{(-1)^{N_\Theta}}{(\gamma_1 - 1)!} (t-\tau)^{\gamma_1 - 1} \tau^{N_\Theta}$ with $N_\Theta = \sum_{l=1}^{\Theta} n_l - j_l$, I_l for $l = 1, \cdots, \Theta$, C_1 and γ_1 have been defined in Lemma 2.2.

Proof: In Lemma 2.2 we set $\bar{m}_l = 0$ for $l = 1, \cdots, \Theta$ and $f = X_N$. As $m_l = \hat{m}_l$ for $l = 1, \cdots, \Theta$, we have

$$\Pi_E\left(\hat{X}_N(s)\right) = \sum_{j_\Theta=0}^{I_\Theta} \cdots \sum_{j_1=0}^{I_1} C_1 \frac{1}{s^{\gamma_1}} \frac{d^{N_\Theta}}{ds^{N_\Theta}} \left(\hat{X}_N(s)\right). \tag{22}$$

Since $\gamma_1 = \sum_{l=1}^{\Theta} m_l + j_l$, condition ($C3$) implies that we have $\gamma_1 \geq \sum_{l=1}^{\Theta} m_l \geq 1$. Using the classical rules of operational calculus, we get

$$\begin{aligned}\mathcal{L}^{-1}\left\{\Pi_E\left(\hat{X}_N(s)\right)\right\}(t) &= \sum_{j_\Theta=0}^{I_\Theta} \cdots \sum_{j_1=0}^{I_1} C_1 \mathcal{L}^{-1}\left\{\frac{1}{s^{\gamma_1}} \frac{d^{N_\Theta}}{ds^{N_\Theta}} \left(\hat{X}_N(s)\right)\right\}(t) \\ &= \sum_{j_\Theta=0}^{I_\Theta} \cdots \sum_{j_1=0}^{I_1} C_1 \frac{(-1)^{N_\Theta}}{(\gamma_1 - 1)!} \int_0^t (t-\tau)^{\gamma_1 - 1} \tau^{N_\Theta} X_N(\tau) d\tau.\end{aligned}$$

∎

Corollary 1 *Let us consider the annihilator Π_E defined by (10) where the set E satisfies conditions ($C1$) and ($C2$) of Proposition 2.1 and condition ($C3$) of Proposition 2.3. Then an estimator of $x^{(n)}(t_0)$ is given by:*

$$\tilde{x}^{(n)}(t_0) = \frac{(r_0 + n)!}{c_n \, \hat{c} \, T^{r_0 + n}} \sum_{i=0}^{L} a_i \int_0^T p_{T,\Theta}(\tau) y(t_0 + \beta_i \tau) d\tau, \tag{23}$$

where $T \in D$, $\hat{c}$ is defined by (12), $r_0 = \sum_{l=1}^{\Theta} m_l + n_l$ and the polynomial $p_{T,\Theta}$ is defined in Proposition 2.3.

By writing $R_N(t) = x(t) - x_N(t)$ and $y(t) = x_N(t) + R_N(t) + \varpi(t)$, we can obtain from (23) that

$$\begin{aligned}\tilde{x}^{(n)}(t_0) =& \frac{(r_0 + n)!}{c_n \, \hat{c} \, T^{r_0 + n}} \sum_{i=0}^{L} a_i \int_0^T p_{T,\Theta}(\tau) \left\{x_N(t_0 + \beta_i \tau) + R_N(t_0 + \beta_i \tau) + \varpi(t_0 + \beta_i \tau)\right\} d\tau \\ =& x^{(n)}(t_0) + e_{R_N}(t_0) + e_{\varpi}(t_0).\end{aligned} \tag{24}$$

So the estimation is corrupted by two sources of errors: $e_{R_N}(t_0)$ the bias term coming from the truncation of the Taylor series expansion of the analytic signal x, and the noise error contribution $e_{\varpi}(t_0)$.

Proof of Corollary 1: We start by applying the annihilator Π_E to the relation (7) and then to go back into the time domain.

Firstly, due to the linearity of $\mathcal{L}^{-1}$, we can apply Proposition 2.3. This yields:

$$\mathcal{L}^{-1}\left\{\Pi_E\left(\hat{X}_N(s)\right)\right\}(t) = \int_0^t p_{t,\Theta}(\tau) X_N(\tau) d\tau. \tag{25}$$

Secondly, by applying the result of Proposition 2.1 to (7), we obtain

$$\Pi_E\left(\sum_{j=0}^{N} c_j s^{-(j+1)} x^{(j)}(t_0)\right) = c_n x^{(n)}(t_0) \frac{\hat{c}}{s^{n+1+r_0}}.$$

As $r_0 = \sum_{l=1}^{\Theta} n_l + m_l \geq \sum_{l=1}^{\Theta} m_l \geq 1$ (condition (C_3)), then we can go back into the time domain and we have:

$$\mathcal{L}^{-1}\left\{\Pi_E\left(\sum_{j=0}^{N} c_j s^{-(j+1)} x^{(j)}(t_0)\right)\right\} = c_n\,\hat{c}\,x^{(n)}(t_0) \frac{t^{r_0+n}}{(r_0+n)!}. \tag{26}$$

Finally, since $c_n \neq 0$, we have

$$x^{(n)}(t_0) = \frac{(r_0+n)!}{c_n\,\hat{c}\,t^{r_0+n}} \int_0^t p_{t,\Theta}(\tau) X_N(\tau) d\tau, \text{ with } t \in D. \tag{27}$$

Recall that $X_N(\tau) = \sum_{i=0}^{L} a_i x_N(t_0 + \beta_i \tau)$, then by substituting $x_N(\cdot)$ by $y(\cdot)$ in (27) it yields

$$\tilde{x}^{(n)}(t_0) = \frac{(r_0+n)!}{c_n\,\hat{c}\,t^{r_0+n}} \sum_{i=0}^{L} a_i \int_0^t p_{t,\Theta}(\tau) y(t_0 + \beta_i \tau) d\tau. \tag{28}$$

Here, the variable t is the length of the estimation time interval. The equation (28) has therefore to be considered for fixed t, say $t = T \in D$. ■

Remark 3 *The corresponding iterated time integrals in* (23) *are low pass filters which attenuate the corrupting noises, which are viewed as highly fluctuating phenomena (see [14] for more details).*

Remark 4 *Note that the noisy function y may not be integrable in* (23). *This expression is only formal. Usually, the observation function y is only known on discrete values. Hence, the integral in* (23) *will be approximated by a numerical integration method.*

Remark 5 *We can consider a new family of linear differential operators defined by* (8), *where $\Pi = \sum_{j=1}^{W} \rho_j \, \Pi_{E_j}$, with $W \in \mathbb{N}^*, \rho_j \in \mathbb{R}^*$ and Π_{E_j} being defined by (10). Moreover, we assume that the set E_j satisfies conditions* $(C1)$, $(C2)$ *and* $(C3)$. *Such affine annihilators help us to estimate $x^{(n)}(t_0)$, which have the following integral form:*

$$\tilde{x}^{(n)}(t_0) = \sum_{j=1}^{W} \frac{(r_0+n)!}{c_n\,\hat{C}_W\,T^{r_0+n}} \sum_{i=0}^{L} a_i \int_0^T p_{T,\Theta,j}(\tau) y(t_0 + \beta_i \tau) d\tau, \tag{29}$$

where $\hat{C}_W = \sum_{j=1}^{W} \rho_j \hat{c}(j)$*,* $\hat{c}(j)$ *and each polynomial* $p_{T,\Theta,j}$ *defined by* (23) *is associated to* Π_{E_j}*.*

We have given conditions on E such that the linear differential operator defined by (10) gives an integral estimator. We will see in the following proposition that we can build some sets E such that the conditions $(C1)$, $(C2)$ and $(C3)$ are satisfied.

Proposition 2.4 *There exists the sets* $E = \{(n_l, m_l)\}_{l=1}^{\Theta}$ *for* $\Theta \geq 3, \Theta \in \mathbb{N}$ *that meet the conditions* $(C1)$ *and* $(C2)$ *given in Proposition* 2.1 *and the condition* $(C3)$ *given in Proposition* 2.3*.*

Proof: We are going to prove firstly that there exists the sets E meeting the conditions $(C1)$ and $(C2)$ given in Proposition 2.1. Each of these sets can give us an annihilator by annihilating all the undesired terms: $c_j s^{-(j+1)} x^{(j)}(t_0)$ in (7) with $j \neq n$ and keeping the term $c_n s^{-(n+1)} x^{(n)}(t_0)$ at the same time. The construction of these sets depends on the way of annihilating the undesired terms, but in any case they can be found. We are going to give such a set.

In order to annihilate the undesired terms, it is necessary to let the degree of s be positive. So, we can choose in particular $n_\Theta = 0$ and $m_\Theta = -n$, we get

$$\frac{1}{s^{m_\Theta}} \frac{d^{n_\Theta}}{ds^{n_\Theta}} \left(\sum_{j=0}^{N} c_j s^{-(j+1)} x^{(j)}(t_0) \right) = c_0 s^{n-1} x(t_0) + \cdots + c_{n-1} x^{(n-1)}(t_0) + c_n s^{-1} x^{(n)}(t_0) + \cdots + c_N s^{-N-1+n} x^{(N)}(t_0). \tag{30}$$

Now, we can annihilate the terms $c_j s^{-(j+1)} x^{(j)}(t_0)$ for $j = 0, \cdots, n-1$ by taking $n_{\Theta-1} = n + k$ with $k \in \mathbb{N}$, we get $0 \leq -(j+1) - (n_\Theta + m_\Theta) \leq n_{\Theta-1}$ for $j = 0, \cdots, n-1$ and

$$\frac{d^{n_{\Theta-1}}}{ds^{n_{\Theta-1}}} \frac{1}{s^{m_\Theta}} \frac{d^{n_\Theta}}{ds^{n_\Theta}} \left(\sum_{j=0}^{N} c_j s^{-(j+1)} x^{(j)}(t_0) \right) = \sum_{j=n}^{N} c_j \frac{(-1)^{n+k}(j+k)!}{(j-n)!} s^{-1-j-k} x^{(j)}(t_0). \tag{31}$$

Then with the same reason, we can annihilate the terms $c_j s^{-(j+1)} x^{(j)}(t_0)$ for $j = n+1, \cdots, N$ by taking $m_{\Theta-1} = -(N+1+k)$ and $n_{\Theta-2} = N - n$, we get $0 \leq -(j+1) - (n_\Theta + m_\Theta + n_{\Theta-1} + m_{\Theta-1}) \leq n_{\Theta-2}$ for $j = n+1, \cdots, N$ and

$$\frac{d^{n_{\Theta-2}}}{ds^{n_{\Theta-2}}} \prod_{i=\Theta-1}^{\Theta} \frac{1}{s^{m_\Theta}} \frac{d^{n_\Theta}}{ds^{n_\Theta}} \left(\sum_{j=0}^{N} c_j s^{-(j+1)} x^{(j)}(t_0) \right) = c_n (-1)^{n+k} (N-n)!(k+n)! x^{(n)}(t_0). \tag{32}$$

Hence, utile here we have found an annihilator meeting the conditions $(C1)$ and $(C2)$. In order to meet the condition $(C3)$, we can choose $m_{\Theta-2} = |m_\Theta| + |m_{\Theta-1}| + \mu + 1 = N + n + k + \mu + 2$ with $\mu \in \mathbb{N}$. Finally, we have found the set $\{(n_l, m_l)\}_{l=\Theta-2}^{\Theta}$ that meets the conditions $(C1)$, $(C2)$ and $(C3)$, and the annihilator associated is

$$\frac{1}{s^{N+n+k+\mu+2}} \frac{d^{N-n}}{ds^{N-n}} s^{N+1+k} \frac{d^{n+k}}{ds^{n+k}} s^n. \tag{33}$$

For $1 \leq l \leq \Theta - 3$, let us take $n_l = m_l = 1$, then the conditions $(C1)$, $(C2)$ and $(C3)$ hold and we have

$$\Pi_E \left(\sum_{j=0}^{N} c_j s^{-(j+1)} x^{(j)}(t_0) \right) = \frac{c_n (-1)^{n+k+\Theta-3} (N-n)!(k+n)! x^{(n)}(t_0)}{s^{N+n+k+\mu+2\Theta-4}} \prod_{i=1}^{\Theta-3} (N + n + k + \mu + 2i). \tag{34}$$

■

Remark 6 *By taking* $\Theta = 3$ *the conditions* $(C1)$*,* $(C2)$ *and* $(C3)$ *are then satisfied.*

Remark 7 *Let us recall the following annihilator used in [35]:*

$$\Pi_E = \frac{1}{s^\nu} \cdot \frac{d^{n+k}}{ds^{n+k}} \cdot \frac{1}{s} \cdot \frac{d^{N-n}}{ds^{N-n}} \cdot s^{N+1}, \quad \nu = N+1+\mu, \mu \in \mathbb{N}, \; k \in \mathbb{N}. \tag{35}$$

where $m_1 = \nu$, $n_1 = n+k$, $m_2 = 1$, $n_2 = N-n$, $m_3 = -(N+1)$, $n_3 = 0$. In particular, we denote it by $\Pi_{k,\mu}^{N,n}$. This set $E = \{(n_l, m_l)\}_{l=1}^{3}$ meets also the conditions $(C1)$, $(C2)$ and $(C3)$. In fact, $\hat{X}_N(s)$ defined by (7) will be a polynomial of degree N if we multiply it by s^{N+1}. Then we annihilate the terms of degree lower than $N-n$ by applying $N-n$ times derivations. For preserving the term including $x^{(n)}(t_0)$, we multiply the remaining polynomial by $1/s$. In order to annihilate the other terms including $x^{(i)}(t_0)$, $i \neq n$, we apply more than n derivations with respect to s. Finally, we multiply by $1/s^\nu$ for obtaining an integral estimator. We will see in Section 3.1 the detailed calculations.

Remark 8 *If we take $N = n$, then $\Pi_{k,\mu}^{N,n}$ reduces to $\Pi_{k,\mu}^{n} = \frac{1}{s^{n+1+\mu}} \cdot \frac{d^{n+k}}{ds^{n+k}} \cdot s^n$. It was shown in [35] that $\Pi_{k,\mu}^{n}$ is an annihilator where in (7) $N = n$ and $\Pi_{k,\mu}^{N,n}$ is an affine annihilator of the following form:*

$$\Pi_{k,\mu}^{N,n} = \sum_{j=0}^{N-n} \left(\sum_{j=i}^{\min(n+k,N-n-i)+i} a_{i,j} \right) \Pi_{k+N-n-j,\mu+j}^{n}, \quad \textit{with } a_{i,j} \in \mathbb{Q}. \tag{36}$$

Inspired by this, let us look at (19) in Lemma 2.2 where Π_E is an annihilator if E meets conditions $(C1)$, $(C2)$ and $(C3)$. We wonder if the $\Pi_{\bar{E}_{\mathbf{j}}}$ are also annihilators similar to the annihilator $\Pi_{k+N-n-j,\mu+j}^{n}$ in (36). By assumption, the $\bar{E}_{\mathbf{j}}$ meet conditions $(C1)$ and $(C2)$. Moreover, from $(C1)$ and $(C2)$, as $\gamma_1 + \sum_{l=1}^{\Theta} \bar{m}_l \geq \sum_{l=1}^{\Theta} m_l$, then condition $(C3)$ holds automatically. Consequently, the annihilator Π_E applied to (7) with $N > n$ will be an affine annihilator of annihilators $\Pi_{\bar{E}_{\mathbf{j}}}$ applied to (7) with $N = n$.

3 Parametric pointwise derivative estimation

We investigate in this section some detailed properties and performances of a class of pointwise derivative estimators. These estimators will be derived from a particular family of annihilators. Moreover, as we shall shortly see, the Jacobi orthogonal polynomials [43], [29] are inherently connected with these estimators.

3.1 Anti-causal estimator and causal estimator

Let us estimate the n^{th} order derivative of x, $0 \leq n \leq N$. To do that, we will use (7) by taking $X(t) = x(t_0 + \beta t)$ with $\beta \in \mathbb{R}^*$ and to which we will apply the annihilator $\Pi_{k,\mu}^{N,n}$ defined by (35). Thus, we will have a family of anti-causal estimators (resp. a family of causal estimators) if $\beta > 0$ (resp. if $\beta < 0$).

Proposition 3.5 *The estimation of the derivative value $x^{(n)}(t_0)$ for any point $t_0 \in I$ is given by:*

$$\tilde{x}_{t_0}^{(n)}(k, \mu, \beta T, N) = \frac{1}{(\beta T)^n} a_{k,\mu,n,N} \sum_{i=0}^{N-n} b_{n,N,i} \, K_i, \tag{37}$$

with

$$K_i = \sum_{j=max(0,k-i)}^{n+k} c_{k,\mu,n,N,j} \int_0^1 p_{k,\mu,N,i,j}(\tau) \, y(\beta T \tau + t_0) d\tau,$$

$$a_{k,\mu,n,N} = (-1)^{n+k} \frac{(\nu+n+k)!}{(n+k)!(N-n)!}, \; b_{n,N,i} = \binom{N-n}{i} \frac{(N+1)!}{(n+i+1)!},$$

$$c_{k,\mu,n,N,j} = \frac{(-1)^{i+j}}{(\nu+k-i-j-1)!} \binom{n+k}{j} \frac{(n+i)!}{(i+j-k)!}, \; p_{k,\mu,N,i,j}(\tau) = (1-\tau)^{\nu+k-i-j-1} \tau^{i+j}.$$

The anti-causal estimator $\tilde{x}_{t_0}^{(n)}(k,\mu,\beta T,N)$ ($\beta>0$) (resp. causal estimator $\tilde{x}_{t_0}^{(n)}(k,\mu,\beta T,N)$ ($\beta<0$)) is obtained by using the integral window $[t_0,t_0+\beta T]\subset I$ (resp. $[t_0+\beta T,t_0]\subset I$), $k\in\mathbb{N}$, $\mu\in\mathbb{N}$, $T>0$ and $\nu=N+1+\mu$.

Proof: Let $X(t)=x(t_0+\beta t)$ with $\beta\in\mathbb{R}^*$ and $t>0$, then (7) becomes

$$\hat{X}_N(s)=\sum_{i=0}^{N}\beta^i s^{-(i+1)}x^{(i)}(t_0), \tag{38}$$

where $\hat{X}_N(s)$ is the Laplace transform of $X_N(t)$. We proceed to annihilate the terms including $x^{(i)}(t_0)$, $i\neq n$ in the right hand side in equation (38) by multiplying the annihilator defined by (35):

$$\Pi_{k,\mu}^{N,n}=\frac{1}{s^\nu}\cdot\frac{d^{n+k}}{ds^{n+k}}\cdot\frac{1}{s}\cdot\frac{d^{N-n}}{ds^{N-n}}\cdot s^{N+1},$$

with $\nu=N+1+\mu$, $\mu\in\mathbb{N}$, $k\in\mathbb{N}$. It reads as

$$\begin{aligned}\Pi_{k,\mu}^{N,n}\left(\hat{X}_N\right)&=\frac{1}{s^\nu}\cdot\frac{d^{n+k}}{ds^{n+k}}\sum_{i=0}^{n}\beta^i\frac{(N-i)!}{(n-i)!}s^{n-i-1}x^{(i)}(t_0)\\&=\frac{\beta^n(N-n)!(-1)^{n+k}(n+k)!}{s^{1+n+k+\nu}}\,x^{(n)}(t_0).\end{aligned}$$

On the other hand, we have

$$\begin{aligned}\Pi_{k,\mu}^{N,n}\left(\hat{X}_N\right)&=\frac{1}{s^\nu}\cdot\frac{d^{n+k}}{ds^{n+k}}\sum_{i=0}^{N-n}\frac{\binom{N-n}{i}(N+1)!}{(n+i+1)!}s^{n+i}(\hat{X}_N)^{(i)}\\&=\sum_{i=0}^{N-n}\frac{\binom{N-n}{i}(N+1)!}{(n+i+1)!}F_i,\end{aligned}$$

with

$$F_i=\sum_{j=\max(0,k-i)}^{n+k}\frac{\binom{n+k}{j}(n+i)!}{(i+j-k-\nu)!}s^{i+j-k}(\hat{X}_N)^{(i+j)}.$$

So, we have

$$\frac{x^{(n)}(t_0)}{s^{\nu+n+k+1}}=\frac{(-1)^{n+k}}{\beta^n(n+k)!(N-n)!}\sum_{i=0}^{N-n}\frac{\binom{N-n}{i}(N+1)!}{(n+i+1)!}\bar{F}_i, \tag{39}$$

with

$$\bar{F}_i=\sum_{j=\max(0,k-i)}^{n+k}\frac{\binom{n+k}{j}(n+i)!}{(i+j-k)!}\frac{(\hat{X}_N)^{(i+j)}}{s^{\nu+k-i-j}}.$$

As $\nu+k-i-j\geq 1$, we can express (39) back into the time domain by using the classical rules of operational calculus and the Cauchy formula for repeated integrals:

$$x^{(n)}(t_0)=\frac{(-1)^{n+k}}{\beta^n T^{\nu+n+k}}\frac{(\nu+n+k)!}{(n+k)!(N-n)!}\sum_{i=0}^{N-n}\sum_{j=\max(0,k-i)}^{n+k}\frac{\binom{N-n}{i}(N+1)!}{(n+i+1)!}A_{i,j}, \tag{40}$$

with

$$A_{i,j}=\frac{(-1)^{i+j}}{(\nu+k-i-j-1)!}\frac{\binom{n+k}{j}(n+i)!}{(i+j-k)!}\int_0^T(T-\tau)^{\nu+k-i-j-1}\tau^{i+j}x_N(\beta\tau+t_0)d\tau.$$

By replacing $x_N(\beta\tau+t_0)$ by the noisy observed signal $y(\beta\tau+t_0)$, a family of estimators can be obtained, which are parameterized by k, μ, T and N. We can achieve the proof by applying the following change of the variable: $\tau\to T\tau$. ■

Remark 9 *Since the set* $E = \{(n_l, m_l)\}_{l=1}^{3}$ *given in* $\Pi_{k,\mu}^{N,n}$ *meets the conditions* $(C1)$ *and* $(C2)$ *from Proposition* 2.1 *and the condition* $(C3)$ *from Proposition* 2.3*, the above proposition can also be obtained by using Corollary* 1*.*

Remark 10 *The algebraic manipulations in the above proof correspond to a linear operator:*

$$\Pi_{k,\mu,T}^{N,n} = I_p \circ R_y \circ \mathcal{L}_T^{-1} \circ \left\{ \frac{(-1)^{n+k}}{\beta^n (n+k)!(N-n)!} \cdot \Pi_{k,\mu}^{N,n} \right\}, \tag{41}$$

where $\mathcal{L}_T^{-1}$ *is the inverse linear operator passing back into time domain at point* $t = T$ *and its restriction in the continuous functions space is injective.The operator* R_y *is the linear operator replacing* $x_N(\tau)$ *by* $y(\tau)$*, and* I_p *is the linear operator taking the change of the variables to* $[0, 1]$*.*

Remark 11 *Let us look at the formula of* $\tilde{x}_{t_0}^{(n)}(k, \mu, \beta T, N)$ *given in* (37)*, if* $\beta > 0$ *(resp. if* $\beta < 0$*), we can take* $X(t) = x(t_0 + t)$ *(resp.* $X(t) = x(t_0 - t)$*) and then have an estimator* $\tilde{x}_{t_0}^{(n)}(k, \mu, \bar{T}, N)$ *that will be equal to* $\tilde{x}_{t_0}^{(n)}(k, \mu, \beta T, N)$*, if we take* $\bar{T} = \beta T$*. So we can assume that* $\beta = \pm 1$*. When* $\beta = 1$*, we denote the anti-causal estimator by* $\tilde{x}_{t_{0+}}^{(n)}(k, \mu, T, N)$*, when* $\beta = -1$*, we denote the causal estimator by* $\tilde{x}_{t_{0-}}^{(n)}(k, \mu, -T, N)$*.*

Remark 12 *If* $N = n$*, we use the simplified notations* $\tilde{x}_{t_{0\pm}}^{(n)}(k, \mu, \pm T)$ *and call these estimators minimal estimators (see [35]). They estimate the* n^{th} *order derivative from an* n^{th} *order truncated Taylor series expansion. Moreover, we denote the bias term errors in these minimal estimators by* $e_{R_n}^{\pm}(t_0; k, \mu, \pm T)$ *and the noise error contributions by* $e_{\varpi}^{\pm}(t_0; k, \mu, \pm T, n)$*.*

3.2 Affine estimator

The estimator defined in Proposition 3.5 can be written as an affine combination of some minimal estimators by using (36). This affine estimator corresponds to a point in the $\mathbb{Q}$-affine hull of the set

$$S_{k,\mu,\pm T,q} = \left\{ \tilde{x}_{t_{0\pm}}^{(n)}(k+q, \mu, \pm T), \ldots, \tilde{x}_{t_{0\pm}}^{(n)}(k, \mu+q, \pm T) \right\} \text{ with } q = N - n. \tag{42}$$

A new affine estimator was introduced in [35], that corresponds to a point in the $\mathbb{R}$-affine hull of the set (42). So, it is clear that any point in this set will represent an n^{th} order derivative estimation of $x(t)$ at point t_0, in some meaningful sense. Characterizing those points which minimize a given distance to $x^{(n)}(t_0)$ is an important question.

Definition 3.6 *Let* n*,* N*,* k*,* $\mu \in \mathbb{N}$ *and a real* $\xi \in [0, 1]$*, then we define an affine anti-causal estimator of the* n^{th} *order derivative of* x *at* t_0 *by*

$$\tilde{x}_{t_{0+}}^{(n)}(k, \mu, T, N, \xi) := \sum_{l=0}^{q} \lambda_l(\xi)\, \tilde{x}_{t_{0+}}^{(n)}(k_l, \mu_l, T), \tag{43}$$

with $[t_0, t_0 + T] \in I$ *and an affine causal estimator of the* n^{th} *order derivative of* x *at* t_0 *by*

$$\tilde{x}_{t_{0-}}^{(n)}(k, \mu, T, N, \xi) := \sum_{l=0}^{q} \lambda_l(\xi)\, \tilde{x}_{t_{0-}}^{(n)}(k_l, \mu_l, -T), \tag{44}$$

with $[t_0 - T, t_0] \in I$*. In these two cases,* $\lambda_l(\xi) \in \mathbb{R}$ *and* $(k_l, \mu_l) = (k + q + l, \mu + l)$*.*

Proposition 3.7 *[35] Let* $\tilde{x}_{t_{0+}}^{(n)}(k, \mu, T, N, \xi)$ *be an affine anti-causal estimator and* $\tilde{x}_{t_{0-}}^{(n)}(k, \mu, -T, N, \xi)$ *be an affine causal estimator. Assume that* $q \leq k + n$ *with* $q = N - n$*, then for any* $\xi \in [0, 1]$*, there exists a unique set of real coordinates* $\lambda_l(\xi) \in \mathbb{R}$*, for* $l = 0, \ldots, q$*, such that*

$$\tilde{x}_{t_{0\pm}}^{(n)}(k, \mu, \pm T, N, \xi) = x_{LS,q}^{(n)}(\pm T\xi) + e_{\varpi}^{\pm}(t_0; k, \mu, \pm T, n, N, \xi), \tag{45}$$

where

$$x_{LS,q}^{(n)}(\pm T\xi) := \sum_{i=0}^{q} \frac{\langle P_i^{k,\mu}(\tau), x^{(n)}(\pm T\tau + t_0)\rangle}{\|P_i^{k,\mu}\|^2} P_i^{k,\mu}(\xi) \tag{46}$$

and

$$e_{\varpi}^{\pm}(t_0; k, \mu, \pm T, n, N, \xi) = \sum_{l=0}^{q} \lambda_l(\xi)\, e_{\varpi}^{\pm}(t_0; k_l, \mu_l, \pm T, n). \tag{47}$$

The $P_i^{k,\mu}$ denotes the Jacobi polynomial and $x_{LS,q}^{(n)}(\pm T\xi)$ denote the least-squares q^{th} order polynomial approximation of $x^{(n)}(\cdot)$ in the interval $[t_0, T + t_0]$ (resp.$[-T + t_0, t_0]$).

Moreover, these coordinates satisfy $\sum_{l=0}^{q} \lambda_l(\xi) = 1.$

Proof: See [35] for the original proof. ■

Remark 13 *It was shown in [35] that the affine causal estimator $\tilde{x}_{t_{0-}}^{(n)}(k, \mu, -T, N, \xi)$ produces a time-delayed estimation of value $\tau = T\xi$. We always take the smallest value of ξ among the roots of the Jacobi polynomial $P_{q+1}^{k,\mu}$, such that the affine causal estimator may be significantly improved by admitting the minimum time delay.*

Remark 14 *The calculation of $\lambda_l(\xi)$ for $l = 0, \cdots, q$ is obtained in [35] by the following formula:*

$$\lambda(\xi) = \Phi^{-1} B^{-1} b_q(\xi), \tag{48}$$

where

$$\lambda(\xi) = \begin{bmatrix} \lambda_0(\xi) \\ \vdots \\ \lambda_q(\xi) \end{bmatrix}, \; b_q(\xi) = \begin{bmatrix} b_{0,q}(\xi) \\ \vdots \\ b_{q,q}(\xi) \end{bmatrix}, \; \Phi = \begin{bmatrix} \Phi_0 & & 0 \\ & \ddots & \\ 0 & & \Phi_q \end{bmatrix}, \tag{49}$$

with $B_{i,j} = \binom{q}{i}\|P_0^{k+2q-(i+j),\mu+(i+j)}\|^2\binom{q}{j}$, for $0 \le i, j \le q$, $b_{l,q}(\xi) = \binom{q}{l}\xi^{q-l}(1-\xi)^l$, $\Phi_l = \frac{\gamma_{k_l,\mu_l,n}}{\binom{q}{l}}$, and $\gamma_{k_l,\mu_l,n} = \frac{(\mu_l+k_l+2n+1)!}{(\mu_l+n)!(k_l+n)!}$, for $0 \le l \le q$. Since $\lambda_l(\xi)$ also depends on the parameters k and μ, we denote it by $\lambda_l(\xi, k, \mu)$, for $l = 0, \ldots, q$.

Remark 15 *[35] When $\xi = 0$, the estimators $\tilde{x}_{t_{0\pm}}^{(n)}(k, \mu, \pm T, N, 0)$ are equal to $\tilde{x}_{t_{0\pm}}^{(n)}(k, \mu, \pm T, N)$ defined in Proposition 3.5. So Proposition 3.7 gives a general anti-causal estimator and a general causal estimator. They can be written as*

$$\tilde{x}_{t_{0\pm}}^{(n)}(k, \mu, \pm T, N, \xi) = \int_0^1 p_{k,\mu,\pm T,n,N,\xi}(\tau)\, y(\pm T\tau + t_0) d\tau, \tag{50}$$

where $p_{k,\mu,\pm T,n,N,\xi}$ are the associated polynomials.

Remark 16 *Since $N = n$ for minimal estimators , $y(\pm T\tau + t_0) = x_n(\pm T\tau + t_0) + R_n(\pm T\tau + t_0) + \varpi(\pm T\tau + t_0)$ where $R_n(\pm T\tau + t_0) = x(\pm T\tau + t_0) - x_n(\pm T\tau + t_0)$. The estimations of $x^{(n)}(t_0)$ are given by (50) and they are corrupted by two sources of error: $e_{R_n}^{\pm}(t_0; k, \mu, \pm T, N, \xi)$ the bias term errors which come from the truncation of the Taylor series expansion of the analytical signal x and the noise error contributions $e_{\varpi}^{\pm}(t_0; k, \mu, \pm T, n, N, \xi)$. They are given by*

$$e_{R_n}^{\pm}(t_0; k, \mu, \pm T, N, \xi) = \int_0^1 p_{k,\mu,\pm T,n,N,\xi}(\tau)\, R_n(\pm T\tau + t_0) d\tau, \tag{51}$$

$$e_{\varpi}^{\pm}(t_0; k, \mu, \pm T, n, N, \xi) = \int_0^1 p_{k,\mu,\pm T,n,N,\xi}(\tau)\; \varpi(\pm T\tau + t_0) d\tau. \tag{52}$$

3.3 Analysis on the bias term error

3.3.1 Analysis on the bias term error for the minimal estimators

In the following proposition, we will give the lower and upper bounds for the bias term errors $e^{\pm}_{R_n}(t_0;k,\mu,\pm T)$ for the minimal estimators $\tilde{x}^{(n)}_{t_{0\pm}}(k,\mu,\pm T)$ defined in Remark 12.

Proposition 3.8 *Let $\tilde{x}^{(n)}_{t_{0\pm}}(k,\mu,\pm T)$ be minimal estimators (see Remark 12), then*

$$\frac{k+n+1}{\mu+k+2n+2}T \inf_{t_0<\theta_+<t_0+T} x^{(n+1)}(\theta_+) \leq e^{+}_{R_n}(t_0;k,\mu,T) \leq \frac{k+n+1}{\mu+k+2n+2}T \sup_{t_0<\theta_+<t_0+T} x^{(n+1)}(\theta_+). \tag{53}$$

$$-\frac{k+n+1}{\mu+k+2n+2}T \sup_{t_0-T<\theta_-<t_0} x^{(n+1)}(\theta_-) \leq e^{-}_{R_n}(t_0;k,\mu,-T) \leq -\frac{k+n+1}{\mu+k+2n+2}T \inf_{t_0-T<\theta_-<t_0} x^{(n+1)}(\theta_-). \tag{54}$$

In order to proof this proposition, an expression of $e^{\pm}_{R_n}(t_0;k,\mu,\pm T)$ is given in the following lemma.

Lemma 3.9 *The bias term errors $e^{\pm}_{R_n}(t_0;k,\mu,\pm T)$ can be written as:*

$$e^{\pm}_{R_n}(t_0;k,\mu,\pm T) = \frac{(k+\mu+2n+1)!}{(k+n)!(\mu+n)!}\int_0^1 (1-\tau)^{\mu+n}\tau^{k+n}\left(x^{(n)}(\pm T\tau+t_0) - x^{(n)}(t_0)\right)d\tau. \tag{55}$$

Proof: Firstly let us calculate the bias term error $e^{+}_{R_n}(t_0;k,\mu,T)$ for the minimal anti-causal estimator $\tilde{x}^{(n)}_{t_{0+}}(k,\mu,T)$. In order to do this, we need to apply the annihilator $\Pi^n_{k,\mu}$ defined in Remark 8 to (38) where $N=n$ and $\beta=1$. So, by coming back into the time domain, we have

$$\mathcal{L}^{-1}\left\{\Pi^n_{k,\mu}\left(\sum_{i=0}^{n} s^{-(i+1)}x^{(i)}(t_0)\right)\right\} = (-1)^{n+k}\frac{(n+k)!}{(k+\mu+2n+1)!}T^{k+\mu+2n+1}x^{(n)}(t_0). \tag{56}$$

On the other hand, by using the Cauchy formula and the classical rules of operational calculus, we obtain

$$\mathcal{L}^{-1}\left\{\Pi^n_{k,\mu}\left(\hat{X}_n\right)\right\} = \frac{(-1)^{n+k}}{(\mu+n)!}\int_0^T (T-\tau)^{n+\mu}\tau^{n+k}x^{(n)}_n(\tau+t_0)d\tau. \tag{57}$$

Then, it yields

$$x^{(n)}(t_0) = \frac{(\mu+k+2n+1)!}{(\mu+n)!(k+n)!}\int_0^T \frac{(T-\tau)^{n+\mu}\tau^{n+k}}{T^{\mu+k+2n+1}}x^{(n)}_n(\tau+t_0)d\tau. \tag{58}$$

Consequently, we obtain the expression of $e^{+}_{R_n}(t_0;k,\mu,T)$:

$$\begin{aligned} e^{+}_{R_n}(t_0;k,\mu,T) &= \frac{(\mu+k+2n+1)!}{(\mu+n)!(k+n)!}\int_0^T \frac{(T-\tau)^{n+\mu}\tau^{n+k}}{T^{\mu+k+2n+1}}\left(x(\tau+t_0)-x_n(\tau+t_0)\right)^{(n)}d\tau \\ &= \frac{(\mu+k+2n+1)!}{(\mu+n)!(k+n)!}\int_0^T \frac{(T-\tau)^{n+\mu}\tau^{n+k}}{T^{\mu+k+2n+1}}\left(x^{(n)}(\tau+t_0)-x^{(n)}(t_0)\right)d\tau. \end{aligned}$$

Finally, by changing of the variables $\tau \to T\tau$, we get

$$e^{+}_{R_n}(t_0;k,\mu,T) = \frac{(k+\mu+2n+1)!}{(k+n)!(\mu+n)!}\int_0^1 (1-\tau)^{\mu+n}\tau^{k+n}\left(x^{(n)}(T\tau+t_0) - x^{(n)}(t_0)\right)d\tau. \tag{59}$$

In order to calculate the bias term error $e^{-}_{R_n}(t_0;k,\mu,-T)$ for the minimal causal estimator $\tilde{x}^{(n)}_{t_{0-}}(k,\mu,-T)$, T is changed by $-T$ in (59). ■

Proof of Proposition 3.8**:** As $x^{(n)}(\pm T\tau+t_0)-x^{(n)}(t_0)$ represent the remainder terms of the Taylor series expansion of $x^{(n)}$, we obtain by applying the Lagrange form that

$$x^{(n)}(T\tau+t_0)-x^{(n)}(t_0)=x^{(n+1)}(\theta_+)T\tau,\ \text{with}\ t_0<\theta_+<t_0+T\tau, \tag{60}$$

$$x^{(n)}(-T\tau+t_0)-x^{(n)}(t_0)=-x^{(n+1)}(\theta_-)T\tau,\ \text{with}\ t_0-T\tau<\theta_-<t_0. \tag{61}$$

So, (55) becomes

$$e^{\pm}_{R_n}(t_0;k,\mu,\pm T)=\pm\frac{(k+\mu+2n+1)!}{(k+n)!(\mu+n)!}T\int_0^1(1-\tau)^{\mu+n}\tau^{k+n+1}x^{(n+1)}(\theta_\pm)d\tau. \tag{62}$$

Since $(1-\tau)^{\mu+n}\tau^{k+n+1}\geq 0$ for any $\tau\in[0,1]$, we obtain

$$(1-\tau)^{\mu+n}\tau^{k+n+1}\inf_{t_0<\theta_+<t_0+T\tau}x^{(n+1)}(\theta_+)\leq(1-\tau)^{\mu+n}\tau^{k+n+1}x^{(n+1)}(\theta_+)\leq(1-\tau)^{\mu+n}\tau^{k+n+1}\sup_{t_0<\theta_+<t_0+T\tau}x^{(n+1)}(\theta_+)$$

and then

$$(1-\tau)^{\mu+n}\tau^{k+n+1}\inf_{t_0<\theta_+<t_0+T}x^{(n+1)}(\theta_+)\leq(1-\tau)^{\mu+n}\tau^{k+n+1}x^{(n+1)}(\theta_+)\leq(1-\tau)^{\mu+n}\tau^{k+n+1}\sup_{t_0<\theta_+<t_0+T}x^{(n+1)}(\theta_+).$$

Moreover, we have $\int_0^1(1-\tau)^{\mu+n}\tau^{k+n+1}d\tau=\frac{(k+n+1)!(\mu+n)!}{(\mu+k+2n+2)!}$. Consequently, we can deduce that

$$\frac{k+n+1}{\mu+k+2n+2}T\inf_{t_0<\theta_+<t_0+T}x^{(n+1)}(\theta_+)\leq e^{+}_{R_n}(t_0;k,\mu,T)\leq\frac{k+n+1}{\mu+k+2n+2}T\sup_{t_0<\theta_+<t_0+T}x^{(n+1)}(\theta_+). \tag{63}$$

Similarly, we obtain

$$-\frac{k+n+1}{\mu+k+2n+2}T\sup_{t_0-T<\theta_-<t_0}x^{(n+1)}(\theta_-)\leq e^{-}_{R_n}(t_0;k,\mu,-T)\leq-\frac{k+n+1}{\mu+k+2n+2}T\inf_{t_0-T<\theta_-<t_0}x^{(n+1)}(\theta_-).$$

■

We can observe that $\frac{k+n+1}{\mu+k+2n+2}T$ is increasing with respect to k and T and is decreasing with respect to μ. So we will take the values of k and T as small as possible and take the value of μ as big as possible in order to reduce the bias term errors for the minimal estimators $\tilde{x}^{(n)}_{t_{0\pm}}(k,\mu,\pm T)$.

Remark 17 *In fact, as it was shown in [35], the value* $\frac{k+n+1}{\mu+k+2n+2}T$ *is the delay in the minimal estimators* $\tilde{x}_{t_{0\pm}}(k,\mu,\pm T)$*. Consequently, by minimizing this delay we may also minimize the bias term error.*

3.3.2 Analysis on the bias term errors for the estimators of the first order derivative

Now, let us consider the bias term errors for the estimators of the first order derivative. By using Proposition 3.8, we can obtain the two following inequalities of the bias term errors for the minimal anti-causal estimator and the minimal causal estimator:

$$\frac{k+2}{\mu+k+4}Tm_l^+\leq e^{+}_{R_1}(k,\mu,T)\leq\frac{k+2}{\mu+k+4}Tm_u^+. \tag{64}$$

$$-\frac{k+2}{\mu+k+4}Tm_l^-\leq e^{-}_{R_1}(k,\mu,-T)\leq-\frac{k+2}{\mu+k+4}Tm_u^-, \tag{65}$$

where $m_l^+=\inf_{t_0<\theta_+<t_0+T}x^{(2)}(\theta_+)$, $m_u^+=\sup_{t_0<\theta_+<t_0+T}x^{(2)}(\theta_+)$, $m_l^-=\sup_{t_0-T<\theta_-<t_0}x^{(2)}(\theta_-)$ and $m_u^-=\inf_{t_0-T<\theta_-<t_0}x^{(2)}(\theta_-)$.

The affine estimators defined in Proposition 3.7 with $N=2$ and $n=1$ are given by:

$$\tilde{x}_{t_{0\pm}}(k,\mu,\pm T,2,\xi)=\lambda_1(\xi,k,\mu)\tilde{x}_{t_{0\pm}}(k,\mu+1,\pm T)+\lambda_0(\xi,k,\mu)\tilde{x}_{t_{0\pm}}(k+1,\mu,\pm T), \tag{66}$$

where $\lambda_1(\xi,k,\mu)=(k+3)-(\mu+k+5)\xi$ and $\lambda_0(\xi,k,\mu)=1-\lambda_1(\xi,k,\mu)$ are calculated according to (48) (see [34] for the detail).

So we can get the bias term errors for the affine estimators

$$e^{\pm}_{R_1}(t_0;k,\mu,\pm T,2,\xi)=\lambda_1(\xi,k,\mu)e^{\pm}_{R_1}(t_0;k,\mu+1,\pm T)+\lambda_0(\xi,k,\mu)e^{\pm}_{R_1}(t_0;k+1,\mu,\pm T). \tag{67}$$

By using (62), we obtain

$$\begin{aligned} e^{\pm}_{R_1}(t_0;k,\mu,\pm T,2,\xi)=&\pm\lambda_1(\xi,k,\mu)\frac{(k+\mu+4)!}{(k+1)!(\mu+2)!}T\int_0^1(1-\tau)^{\mu+2}\tau^{k+2}x^{(2)}(\theta_\pm)d\tau\\ &\pm\lambda_0(\xi,k,\mu)\frac{(k+\mu+4)!}{(k+2)!(\mu+1)!}T\int_0^1(1-\tau)^{\mu+1}\tau^{k+3}x^{(2)}(\theta_\pm)d\tau\\ =&\pm\frac{(k+\mu+4)!}{(k+2)!(\mu+2)!}T\int_0^1 p_{k,\mu,\xi}(\tau)(1-\tau)^{\mu+1}\tau^{k+2}x^{(2)}(\theta_\pm)d\tau, \end{aligned} \tag{68}$$

where $p_{k,\mu,\xi}(\tau)=(k+2)\left((k+3)-(\mu+k+5)\xi\right)+\left[(\mu+2)-(k+\mu+4)\left((k+3)-(\mu+k+5)\xi\right)\right]\tau$.

Remark 18 *If we take $\xi=\frac{k+2}{\mu+k+5}$, then we will obtain $e^{\pm}_{R_1}(t_0;k,\mu,\pm T,2,\xi)=e^{\pm}_{R_1}(t_0;k,\mu+1,\pm T)$ and $\tilde{\dot{x}}_{t_{0\pm}}(k,\mu,\pm T,2,\xi)=\tilde{\dot{x}}_{t_{0\pm}}(k,\mu+1,\pm T)$.*

We are going to give the lower and upper bounds for $e^{\pm}_{R_1}(t_0;k,\mu,\pm T,2,\xi)$ in the following proposition. We denote the root of $p_{k,\mu,\xi}(\tau)$ by $\tau_0=\frac{-(k+2)((k+3)-(\mu+k+5)\xi)}{(\mu+2)-(k+\mu+4)((k+3)-(\mu+k+5)\xi)}$ and $I_\alpha=\int_0^\alpha p_{k,\mu,\xi}(\tau)(1-\tau)^{\mu+1}\tau^{k+2}d\tau$ where $0<\alpha\le 1$. Moreover, we denote $m^+_{\tau_0,l}=\inf\limits_{t_0<\theta_+<t_0+T\tau_0}x^{(2)}(\theta_+)$, $m^+_{\tau_0,u}=\sup\limits_{t_0<\theta_+<t_0+T\tau_0}x^{(2)}(\theta_+)$, $m^-_{\tau_0,l}=\sup\limits_{t_0-T\tau_0<\theta_-<t_0}x^{(2)}(\theta_-)$ and $m^-_{\tau_0,u}=\inf\limits_{t_0-T\tau_0<\theta_-<t_0}x^{(2)}(\theta_-)$.

Proposition 3.10 *Let $\tilde{\dot{x}}_{t_{0\pm}}(k,\mu,\pm T,2,\xi)$ be affine estimators, then*

$$\pm\frac{(k+\mu+4)!}{(k+2)!(\mu+2)!}TM^{\pm}_l\le e^{\pm}_{R_1}(t_0;k,\mu,\pm T,2,\xi)\le\pm\frac{(k+\mu+4)!}{(k+2)!(\mu+2)!}TM^{\pm}_u, \tag{69}$$

where

$$M^{\pm}_l=\begin{cases} I_{\tau_0}m^{\pm}_{\tau_0,l}+(I_1-I_{\tau_0})\,m^{\pm}_u, & \text{if}\quad 0\le\xi\le\frac{k+2}{k+\mu+5},\\ I_1m^{\pm}_l, & \text{if}\quad \frac{k+2}{k+\mu+5}\le\xi\le\frac{k+3}{k+\mu+5},\\ I_{\tau_0}m^{\pm}_{\tau_0,u}+(I_1-I_{\tau_0})\,m^{\pm}_l, & \text{if}\quad \frac{k+3}{k+\mu+5}\le\xi\le 1, \end{cases} \tag{70}$$

and

$$M^{\pm}_u=\begin{cases} I_{\tau_0}m^{\pm}_{\tau_0,u}+(I_1-I_{\tau_0})\,m^{\pm}_l, & \text{if}\quad 0\le\xi\le\frac{k+2}{k+\mu+5},\\ I_1m^{\pm}_u, & \text{if}\quad \frac{k+2}{k+\mu+5}\le\xi\le\frac{k+3}{k+\mu+5},\\ I_{\tau_0}m^{\pm}_{\tau_0,l}+(I_1-I_{\tau_0})\,m^{\pm}_u, & \text{if}\quad \frac{k+3}{k+\mu+5}\le\xi\le 1. \end{cases} \tag{71}$$

Proof: From (68) we have $p_{k,\mu,\xi}(\tau)=(k+2)\lambda+((\mu+2)-(k+\mu+4)\lambda)\,\tau$, where $\lambda=(k+3)-(\mu+k+5)\xi$. We are going to study the sign of this polynomial in the three following cases.

Firstly, if $(\mu+2)-(k+\mu+4)\lambda=0$ *i.e.* $\lambda=\frac{\mu+2}{k+\mu+4}$, then $p_{k,\mu,\xi}(\tau)\ge 0$ for $0\le\tau\le 1$. Secondly, if $(\mu+2)-(k+\mu+4)\lambda>0$ *i.e.* $\lambda<\frac{\mu+2}{k+\mu+4}$, then we need to take

$$\tau\ge\tau_0=\frac{-(k+2)\lambda}{(\mu+2)-(k+\mu+4)\lambda} \tag{72}$$

so as to $p_{k,\mu,\xi}(\tau) \geq 0$. However, since $k+2>0$ and $((\mu+2)-(k+\mu+4)\lambda)>0$, then we have $\frac{-(k+2)\lambda}{(\mu+2)-(k+\mu+4)\lambda} \leq 0$ for $\lambda \geq 0$. Consequently, if $0 \leq \lambda < \frac{\mu+2}{k+\mu+4}$, then $p_{k,\mu}(\tau,\xi) \geq 0$ for $0 \leq \tau \leq 1$. Moreover, we have $0<\tau_0<1$ for any $\lambda<0$. So we can deduce for $\lambda<0$ that $p_{k,\mu,\xi}(\tau) \geq 0$ for $\tau_0 \leq \tau \leq 1$ and $p_{k,\mu,\xi}(\tau) \leq 0$ for $0 \leq \tau \leq \tau_0$.

Thirdly, if $(\mu+2)-(k+\mu+4)\lambda<0$ *i.e.* $\lambda > \frac{\mu+2}{k+\mu+4}$, then we need to take

$$\tau \leq \tau_0 = \frac{-(k+2)\lambda}{(\mu+2)-(k+\mu+4)\lambda} \tag{73}$$

so as to $p_{k,\mu,\xi}(\tau) \geq 0$. However, if $\frac{\mu+2}{k+\mu+4} < \lambda \leq 1$, then we obtain $\tau_0 \geq 1$ so that $p_{k,\mu,\xi}(\tau) \geq 0$ for $0 \leq \tau \leq 1$. Moreover, if $\lambda>1$, then we have $0<\tau_0<1$. So we can deduce for $\lambda>1$ that $p_{k,\mu,\xi}(\tau) \geq 0$ for $0 \leq \tau \leq \tau_0$ and $p_{k,\mu,\xi}(\tau) \leq 0$ for $\tau_0 \leq \tau \leq 1$.

Let us recall that $\lambda=(k+3)-(\mu+k+5)\xi$ and $0 \leq \xi \leq 1$. Consequently, we can conclude that

$$\begin{cases} p_{k,\mu,\xi}(\tau) \geq 0 \text{ for } 0 \leq \tau \leq \tau_0 \text{ and } p_{k,\mu,\xi}(\tau) \leq 0 \text{ for } \tau_0 \leq \tau \leq 1 & \text{if } \ 0 \leq \xi \leq \frac{k+2}{k+\mu+5}, \\ p_{k,\mu,\xi}(\tau) \geq 0 \text{ for } 0 \leq \tau \leq 1 & \text{if } \ \frac{k+2}{k+\mu+5} \leq \xi \leq \frac{k+3}{k+\mu+5}, \\ p_{k,\mu,\xi}(\tau) \leq 0 \text{ for } 0 \leq \tau \leq \tau_0 \text{ and } p_{k,\mu,\xi}(\tau) \geq 0 \text{ for } \tau_0 \leq \tau \leq 1 & \text{if } \ \frac{k+3}{k+\mu+5} \leq \xi \leq 1, \end{cases} \tag{74}$$

Now, consider the case where $0 \leq \xi \leq \frac{k+2}{k+\mu+5}$. Then we obtain that for $0 \leq \tau \leq \tau_0$

$$\pm p_{k,\mu,\xi}(\tau)(1-\tau)^{\mu+1}\tau^{k+2}m^{\pm}_{\tau_0,l} \leq \pm p_{k,\mu,\xi}(\tau)(1-\tau)^{\mu+1}\tau^{k+2}x^{(2)}(\theta_{\pm}) \leq \pm p_{k,\mu,\xi}(\tau)(1-\tau)^{\mu+1}\tau^{k+2}m^{\pm}_{\tau_0,u}, \tag{75}$$

and for $\tau_0 \leq \tau \leq 1$

$$\pm p_{k,\mu,\xi}(\tau)(1-\tau)^{\mu+1}\tau^{k+2}m^{\pm}_{u} \leq \pm p_{k,\mu,\xi}(\tau)(1-\tau)^{\mu+1}\tau^{k+2}x^{(2)}(\theta_{\pm}) \leq \pm p_{k,\mu,\xi}(\tau)(1-\tau)^{\mu+1}\tau^{k+2}m^{\pm}_{l}. \tag{76}$$

By integrating (75) on the interval $[0,\tau_0]$, we obtain

$$\pm I_{\tau_0} m^{\pm}_{\tau_0,l} \leq \pm \int_0^{\tau_0} p_{k,\mu,\xi}(\tau)(1-\tau)^{\mu+1}\tau^{k+2}x^{(2)}(\theta_{\pm})d\tau \leq \pm I_{\tau_0} m^{\pm}_{\tau_0,u}, \tag{77}$$

and by integrating (76) on the interval $[\tau_0,1]$, we obtain

$$\pm(I_1-I_{\tau_0})m^{\pm}_{u} \leq \pm \int_{\tau_0}^{1} p_{k,\mu,\xi}(\tau)(1-\tau)^{\mu+1}\tau^{k+2}x^{(2)}(\theta_{\pm})d\tau \leq \pm(I_1-I_{\tau_0})m^{\pm}_{l}. \tag{78}$$

Consequently, we obtain two bounds for the bias term errors for the affine estimators when $0 \leq \xi \leq \frac{k+2}{k+\mu+5}$:

$$\pm\frac{(k+\mu+4)!}{(k+2)!(\mu+2)!}TM^{\pm}_{l} \leq e^{\pm}_{R_1}(t_0;k,\mu,\pm T,2,\xi) \leq \pm\frac{(k+\mu+4)!}{(k+2)!(\mu+2)!}TM^{\pm}_{u}, \tag{79}$$

where $M^{\pm}_{l} = I_{\tau_0} m_{\tau_0,l} l^{\pm} + (I_1-I_{\tau_0})\, m^{\pm}_{u}$ and $M^{\pm}_{u} = I_{\tau_0} m^{\pm}_{\tau_0,u} + (I_1-I_{\tau_0})\, m^{\pm}_{l}$.

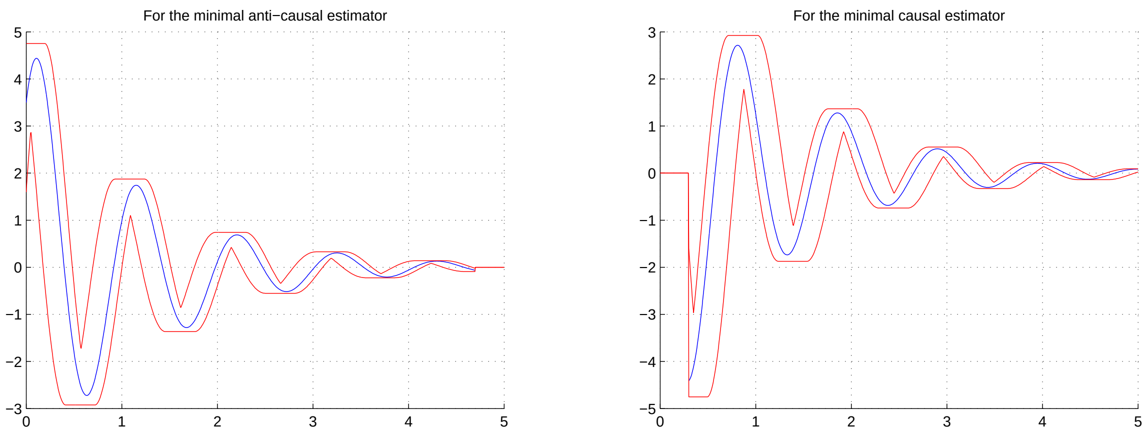
The two other cases for ξ are obtained similarly. ■

In the following example we compare the bias term errors with their bounds. We assume that $x(t)=\tanh(t-1)+\exp(-t/1.2)\sin(6t+\pi)$ for $0<t<5$. In Figure 1 and Figure 2 where $k=\mu=0$ and $T=0.3$, we can see that the bounds for the bias term errors are near optimal.

3.4 Central estimator

The bias term error plays a key role in the total error. So we consider now the two following functions in order to reduce the number of terms in the remainder term of the Taylor series expansion. For any $t_0 \in I$ and for $t>0$, we consider

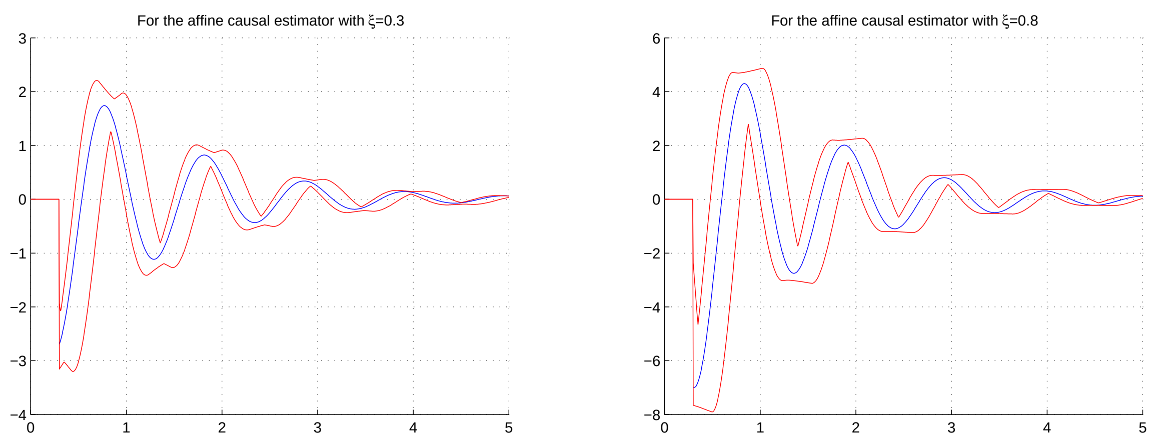
$$X^{-}(t) = \frac{1}{2}(x(t+t_0)-x(-t+t_0)) \ \text{ and } \ X^{+}(t) = \frac{1}{2}(x(t+t_0)+x(-t+t_0)).$$



(a) $e^{+}_{R_1}(k,\mu,T)$ and its bounds

(b) $e^{-}_{R_1}(k,\mu,-T)$ and its bounds

Figure 1: Comparisons between the bias term errors for the minimal estimators and their bounds



(a) $e^{-}_{R_1}(k,\mu,-T,2,\xi)$ and its bounds with $\xi=0.3$

(b) $e^{-}_{R_1}(k,\mu,-T,2,\xi)$ and its bounds with $\xi=0.8$

Figure 2: Comparisons between the bias term errors for the affine causal estimator and their bounds

Take the N^{th} order truncated Taylor series expansion of x. It yields

$$X^{-}_N(t)=\frac{1}{2}\left(x_N(t+t_0)-x_N(-t+t_0)\right) \quad \text{and} \quad X^{+}_N(t)=\frac{1}{2}\left(x_N(t+t_0)+x_N(-t+t_0)\right).$$

If $N\in\mathbb{N}$ is odd, then

$$X^{-}_N(t)=\sum_{i=0}^{\frac{N-1}{2}}\frac{t^{2i+1}}{(2i+1)!}x^{(2i+1)}(t_0) \quad \text{and} \quad X^{+}_N(t)=\sum_{i=0}^{\frac{N-1}{2}}\frac{t^{2i}}{(2i)!}x^{(2i)}(t_0). \tag{80}$$

If $N\in\mathbb{N}$ is even, then

$$X^{-}_N(t)=\sum_{i=0}^{\frac{N}{2}-1}\frac{t^{2i+1}}{(2i+1)!}x^{(2i+1)}(t_0) \quad \text{and} \quad X^{+}_N(t)=\sum_{i=0}^{\frac{N}{2}}\frac{t^{2i}}{(2i)!}x^{(2i)}(t_0). \tag{81}$$

So, X_N^- contains the values of the odd order derivatives at t_0 and X_N^+ contains the values of the even order derivatives at t_0. In the following proposition, we are going to give a new family of estimators based on an affine combination of anti-causal estimators and causal estimators. We call them central estimators.

Proposition 3.11 *The central estimator $\tilde{x}_{t_0}^{(n)}(k,\mu,T,N)$ of the derivative value $x^{(n)}$ at any point $t_0 \in I$ such that $[t_0-T,t_0+T] \subset I$ with $T>0$ is given by:*

$n \setminus N$	even	odd
even	$\frac{1}{2}\left(\tilde{x}_{t_0+}^{(n)}(k,\mu,T,N)+\tilde{x}_{t_0-}^{(n)}(k,\mu,-T,N)\right)$	$\frac{1}{2}\left(\tilde{x}_{t_0+}^{(n)}(k,\mu,T,N-1)+\tilde{x}_{t_0-}^{(n)}(k,\mu,-T,N-1)\right)$
odd	$\frac{1}{2}\left(\tilde{x}_{t_0+}^{(n)}(k-1,\mu,T,N-1)+\tilde{x}_{t_0-}^{(n)}(k-1,\mu,-T,N-1)\right)$	$\frac{1}{2}\left(\tilde{x}_{t_0+}^{(n)}(k-1,\mu,T,N)+\tilde{x}_{t_0-}^{(n)}(k-1,\mu,-T,N)\right)$

where $k \in \mathbb{N}$ and $\mu \in \mathbb{N}$.

Remark 19 *When $k=0$ these expressions still have sense, since in (37) the terms of $\tilde{x}_{t_{0\pm}}^{(n)}(-1,\mu,\pm T,N)$ are defined.*

Proof: The proof of this proposition is similar to the one of Proposition 3.5. For this, we use the differential operators defined by (41) of the following form: $\Pi_{k-1,\mu,T}^{N,n}$, if n and N are odd, $\Pi_{k-1,\mu,T}^{N-1,n}$, if n is odd and N is even, $\Pi_{k,\mu,T}^{N-1,n}$, if n is even and N is odd, $\Pi_{k,\mu,T}^{N,n}$, if n and N are even.

For instance, if n and N are odd, then by applying the operator $\Pi_{k-1,\mu,T}^{N,n}$ to $X_N^-(t)$ which is defined by (80), we obtain

$$\begin{aligned}&\tilde{x}_{t_0}^{(n)}(k,\mu,T,N)\\ =&\frac{1}{(T)^n}a_{k-1,\mu,n,N}\sum_{i=0}^{N-n}b_{n,N,i}\sum_{j=\max(0,k-1-i)}^{n+k-1}c_{k-1,\mu,n,N,j}\int_0^1 p_{k-1,\mu,N,i,j}(\tau)\frac{1}{2}\left(y(T\tau+t_0)-y(-T\tau+t_0)\right)d\tau,\end{aligned}$$

where $a_{k-1,\mu,n,N}$, $b_{n,N,i}$, $c_{k-1,\mu,n,N,j}$ and $p_{k-1,\mu,N,i,j}(\tau)$ are defined in Proposition 3.5.

According to (37), it yields

$$\begin{aligned}\tilde{x}_{t_0}^{(n)}(k,\mu,T,N)&=\frac{1}{2}\left(\tilde{x}_{t_0+}^{(n)}(k-1,\mu,T,N)-(-1)^n\tilde{x}_{t_0-}^{(n)}(k-1,\mu,-T,N)\right)\\ &=\frac{1}{2}\left(\tilde{x}_{t_0+}^{(n)}(k-1,\mu,T,N)+\tilde{x}_{t_0-}^{(n)}(k-1,\mu,-T,N)\right),\quad \text{since } n \text{ is odd.}\end{aligned}$$

The calculations in the other cases are similar. ■

Remark 20 *For odd n and N, we denote by $e_\varpi(t_0;k,\mu,T,n,N)$ the noise error contribution for the central estimator. Then we have*

$$e_\varpi(t_0;k,\mu,T,n,N)=\frac{1}{2}\left(e_\varpi^+(t_0;k-1,\mu,T,n,N)+e_\varpi^-(t_0;k-1,\mu,-T,n,N)\right).$$

When n and N are odd, we can introduce a new family of affine estimators which are an affine combination of these central estimators. As we can see in the following proposition that they are issued from the sum of the projection of $[X^-]^{(n)}(\cdot)$ into the Jacobi polynomials basis and a term due to the noise.

Definition 3.12 *Let n, N be two odd integers, $\mu \in \mathbb{N}$, $k \in \mathbb{N}^*$ and a real $\xi \in [0,1]$, then we define an affine central estimator of the n^{th} order derivative of $x(t_0)$ by*

$$\tilde{x}_{t_0}^{(n)}(k,\mu,T,N,\xi):=\sum_{l=0}^{q}\bar{\lambda}_l(\xi)\,\tilde{x}_{t_0}^{(n)}(k_l,\mu_l,T), \tag{82}$$

where $[t_0-T,t_0+T] \in I$, $\bar{\lambda}_l(\xi) \in \mathbb{R}$ and $(k_l,\mu_l)=(k+q+l,\mu+l)$.

Proposition 3.13 *Let $\tilde{x}_{t_0}^{(n)}(k,\mu,T,N,\xi)$ be an affine central estimator where n and N are odd. Assume that $q \leq k-1+n$ with $q=N-n$, then for any $\xi \in [0,1]$, there exists a unique set of real coordinates $\bar{\lambda}_l(\xi) \in \mathbb{R}$, for $l=0,\ldots,q$, such that*

$$\tilde{x}_{t_0}^{(n)}(k,\mu,T,N,\xi) = [X^-]_{ls,q}^{(n)}(T\xi) + e_{\varpi}(t_0;k,\mu,T,n,N,\xi), \tag{83}$$

where

$$[X^-]_{ls,q}^{(n)}(T\xi) := \sum_{i=0}^{q} \frac{\langle P_i^{k-1,\mu}(\tau), [X^-]^{(n)}(T\tau)\rangle}{\|P_i^{k-1,\mu}\|^2} P_i^{k-1,\mu}(\xi) \tag{84}$$

and

$$e_{\varpi}(t_0;k,\mu,T,n,N,\xi) = \frac{1}{2}\left(e_{\varpi}^{+}(t_0;k-1,\mu,T,n,N,\xi) + e_{\varpi}^{-}(t_0;k-1,\mu,-T,n,N,\xi)\right). \tag{85}$$

$P_i^{k-1,\mu}$ denotes the Jacobi polynomials and $[X^-]_{ls,q}^{(n)}(T\xi)$ denotes the least-squares q^{th} order polynomial approximation of $[X^-]^{(n)}(\cdot)$ in the interval $[t_0, T+t_0]$.

Proof: Using Proposition 3.11, $\bar{\lambda}_l(\xi) \in \mathbb{R}$, for $l=0,\ldots,q$, we have

$$\begin{aligned}\sum_{l=0}^{q} \bar{\lambda}_l(\xi)\, \tilde{x}_{t_0}^{(n)}(k_l,\mu_l,T) &= \sum_{l=0}^{q} \bar{\lambda}_l(\xi)\, \frac{1}{2}\left(\tilde{x}_{t_{0+}}^{(n)}(k_l-1,\mu_l,T) + \tilde{x}_{t_{0-}}^{(n)}(k_l-1,\mu_l,-T)\right)\\ &= \frac{1}{2}\sum_{l=0}^{q} \bar{\lambda}_l(\xi)\, \tilde{x}_{t_{0+}}^{(n)}(k_l-1,\mu_l,T) + \frac{1}{2}\sum_{l=0}^{q} \bar{\lambda}_l(\xi)\, \tilde{x}_{t_{0-}}^{(n)}(k_l-1,\mu_l,-T).\end{aligned}$$

Since $q \leq k-1+n$, we can choose $\bar{\lambda}_l(\xi)$ defined in Proposition 3.7. By using the definitions of affine central estimator, affine anti-causal estimator and affine causal estimator we have

$$\begin{aligned}\tilde{x}_{t_0}^{(n)}(k,\mu,T,N,\xi) &= \frac{1}{2}\left(\tilde{x}_{t_{0+}}^{(n)}(k-1,\mu,T,N,\xi) + \tilde{x}_{t_{0-}}^{(n)}(k-1,\mu,-T,N,\xi)\right)\\ &= \sum_{i=0}^{q} \frac{\langle P_i^{k-1,\mu}(\tau), 1/2\left(x^{(n)}(T\tau+t_0) + x^{(n)}(-T\tau+t_0)\right)\rangle}{\|P_i^{k-1,\mu}\|^2} P_i^{k-1,\mu}(\xi)\\ &\quad + \frac{1}{2}\left(e_{\varpi}^{+}(t_0;k-1,\mu,T,n,N,\xi) + e_{\varpi}^{-}(t_0;k-1,\mu,-T,n,N,\xi)\right).\end{aligned}$$

By denoting $e_{\varpi}(t_0;k,\mu,T,n,N,\xi) = \frac{1}{2}\left(e_{\varpi}^{+}(t_0;k-1,\mu,T,n,N,\xi) + e_{\varpi}^{-}(t_0;k-1,\mu,-T,n,N,\xi)\right)$, the proof can be achieved by using the fact that $[X^-]^{(n)}(T\tau) = \frac{1}{2}\left(x^{(n)}(T\tau+t_0) + x^{(n)}(-T\tau+t_0)\right)$, since n is odd. ■

Remark 21 *As in the above proof, we have another expression for $\tilde{x}_{t_0}^{(n)}(k,\mu,T,N,\xi)$:*

$$\tilde{x}_{t_0}^{(n)}(k,\mu,T,N,\xi) = \frac{1}{2}\left(\tilde{x}_{t_{0+}}^{(n)}(k-1,\mu,T,N,\xi) + \tilde{x}_{t_{0-}}^{(n)}(k-1,\mu,-T,N,\xi)\right). \tag{86}$$

Moreover, according to Remark 48, $\bar{\lambda}_l(\xi)$ also depends on the parameters k and μ. We denote it by $\bar{\lambda}_l(\xi,k,\mu)$ and we have $\bar{\lambda}_l(\xi,k,\mu) = \lambda_l(\xi,k-1,\mu)$, where the $\lambda_l(\xi,k-1,\mu)$ is given in Proposition 3.7, for $l=0,\ldots,q$.

4 Analysis of the noise error contribution in discrete case

Assume now that $y(t_i) = x(t_i) + \varpi(t_i)$ is a noisy measurement of x in a discrete case with an equidistant sampling period T_s, where the noise $\varpi(t_i)$ is a sequence of independent random variables with the same expected value and the same variance. The estimate of the n^{th} order derivative of the signal is given by (23) which can be rewritten as follows:

$$\tilde{x}^{(n)}(t_0) = \frac{(r_0+n)!}{c_n \hat{c} T^{r_0+n-1}} \int_0^1 p_{T,\Theta}(T\tau) \sum_{\gamma=0}^{L} a_\gamma y(t_0 + \beta_\gamma T\tau)\, d\tau. \tag{87}$$

Since $y(\cdot)$ is a discrete measurement, we need to use a numerical integration method in order to approximate the integral value in (87).

4.1 Analysis with a known noise

Let f be a continuous function defined on a bounded interval $f : [0,1] \to \mathbb{R}$. By applying a quadrature formula, the numerical integration approximations of the integral $I = \int_0^1 f(x)\, dx$ are given by:

$$I_m = \sum_{j=0}^{M-1} h \sum_{i=1}^{l} b_i f(x_{(l-1)j} + c_i h), \tag{88}$$

where M and l take values in $\mathbb{N}^*$. As $m = M(l-1)$, we deduce that $h = \frac{1}{M}$ and $x_i = \frac{i}{m}$ for $i = 0, \cdots, m$. The nodes c_i are equal to $c_i = \frac{i-1}{l-1}$ and b_i are the weights of the different classical numerical methods used. For instance, for $l \le 7$, b_i are given in [27]. By applying a numerical integration scheme to (87), it yields

$$I_m = h \frac{(r_0+n)!}{c_n \hat{c} T^{r_0+n-1}} \sum_{j=0}^{M-1} \sum_{i=1}^{l} b_i p_{T,\Theta}(T\tau_{i,j}) \sum_{\gamma=0}^{L} a_\gamma y(t_0 + \beta_\gamma T\tau_{i,j}), \tag{89}$$

where $\tau_{i,j} = \tau_{(l-1)j} + c_i h$. Consequently, the noise error contribution $e_\varpi(t_0)$ can be given, in the discrete case, by

$$e_{\varpi,m}(t_0) = h \frac{(r_0+n)!}{c_n \hat{c} T^{r_0+n-1}} \sum_{j=0}^{M-1} \sum_{i=1}^{l} b_i\, p_{T,\Theta}(T\tau_{i,j}) \sum_{\gamma=0}^{L} a_\gamma\, \varpi(t_0 + \beta_\gamma T\tau_{i,j}). \tag{90}$$

Remark 22 *If we apply a numerical integration method, then an error occurs from this method. Then* (24) *becomes*

$$\tilde{x}_m^{(n)}(t_0) = x^{(n)}(t_0) + e_m(t_0) + e_{R_N,m}(t_0) + e_{\varpi,m}(t_0), \tag{91}$$

where $\tilde{x}_m^{(n)}(t_0)$ is the estimation of the n^{th} order derivative of x in the discrete case, $e_{R_N,m}(t_0)$ is the bias term error and $e_m(t_0) = x_m^{(n)}(t_0) - x^{(n)}(t_0)$ is the numerical integration error.

As $e_{\varpi,m}(t_0)$ is a finite sum of independent random variables with the same expected value and the same variance, we can compute the expected value and the variance of $e_{\varpi,m}(t_0)$. To do this, let $\Gamma = \{(\gamma, i, j);\ 0 \le \gamma \le L, 1 \le i \le l, 0 \le j \le M-1\}$ and $\Gamma_0 = \{(\gamma, i, j);\ (\gamma, i, j) \in \Gamma \text{ such that } \exists (\gamma_1, i_1, j_1) \in \Gamma, \beta_{\gamma_1}\tau_{i_1,j_1} = \beta_\gamma \tau_{i,j}\}$ be two sets of discretized indexes. $\Gamma \setminus \Gamma_0$ denotes the set of indexes of all distinct values of y in (89). From all the values $\beta_\gamma \tau_{i,j}$ where $(\gamma, i, j) \in \Gamma_0$, we define the set Q by setting these values in an increasing order *i.e.* $Q := \{q_g, 1 \le g \le L_0\}$ where $L_0 = card(Q)$. Then, we can give the expected value and the variance of $e_{\varpi,m}(t_0)$ in the following proposition.

Proposition 4.14 *Let $\varpi(t_i)$ be independent random variables with the same expected value $\bar{\alpha} = E[\varpi]$ and the same variance $\bar{\beta} = var[\varpi]$. The expected value of $e_{\varpi,m}(t_0)$ is given by*

$$E[e_{\varpi,m}(t_0)] = \frac{\bar{\alpha}\, h\, (r_0+n)!}{c_n \hat{c} T^{r_0+n-1}} \left(\sum_{\gamma=0}^{L} a_\gamma \right) \sum_{j=0}^{M-1} \sum_{i=1}^{l} b_i\, p_{T,\Theta}(T\tau_{i,j}), \tag{92}$$

and the variance of $e_{\varpi,m}(t_0)$ *is given by*

$$var[e_{\varpi,m}(t_0)] = \frac{\bar{\beta}(h\,(r_0+n)!)^2}{(c_n\hat{c}\,T^{r_0+n-1})^2}(A+B), \tag{93}$$

where $A = \sum_{(\gamma,i,j)\in\Gamma/\Gamma_0} a_\gamma^2\, b_i^2 p_{T,\Theta}^2(T\tau_{i,j})$ *and* $B = \sum_{g=1}^{L_0}\left(\sum_{\substack{q_g=\beta_\gamma\tau_{i,j},\\(\gamma,i,j)\in\Gamma_0}} a_\gamma\, b_i\, p_{T,\Theta}(T\tau_{i,j})\right)^2.$

Proof: It is clear that

$$\begin{aligned}
E[e_{\varpi,m}(t_0)] &= h\frac{(r_0+n)!}{c_n\hat{c}\,T^{r_0+n-1}}\sum_{j=0}^{M-1}\sum_{i=1}^{l} b_i p_{T,\Theta}(T\tau_{i,j})\sum_{\gamma=0}^{L} a_\gamma E[\varpi(t_0+\beta_\gamma T\tau_{i,j})] \\
&= \frac{\bar{\alpha}\,h\,(r_0+n)!}{c_n\hat{c}\,T^{r_0+n-1}}\left(\sum_{\gamma=0}^{L} a_\gamma\right)\sum_{j=0}^{M-1}\sum_{i=1}^{l} b_i\, p_{T,\Theta}(T\tau_{i,j}).
\end{aligned}$$

In order to calculate the variance of $e_{\varpi,m}(t_0)$, it is enough to write $e_{\varpi,m}(t_0)$ with independent terms. According to (90), we have

$$\begin{aligned}
e_{\varpi,m}(t_0) &= h\frac{(r_0+n)!}{c_n\hat{c}\,T^{r_0+n-1}}\left[\sum_{j=0}^{M-1}\sum_{i=1}^{l}\sum_{\gamma=0}^{L} b_i\, p_{T,\Theta}(T\tau_{i,j})\, a_\gamma\varpi(t_0+\beta_\gamma T\tau_{i,j})\right] \\
&= h\frac{(r_0+n)!}{c_n\hat{c}\,T^{r_0+n-1}}\sum_{(\gamma,i,j)\in\Gamma/\Gamma_0} b_i\, p_{T,\Theta}(T\tau_{i,j})\, a_\gamma\varpi(t_0+\beta_\gamma T\tau_{i,j}) \\
&+ h\frac{(r_0+n)!}{c_n\hat{c}\,T^{r_0+n-1}}\sum_{g=1}^{L_0}\sum_{\substack{q_g=\beta_\gamma\tau_{i,j},\\(\gamma,i,j)\in\Gamma_0}} b_i\, p_{T,\Theta}(T\tau_{i,j})\, a_\gamma\ \varpi(t_0+Tq_g).
\end{aligned}$$

Finally, by applying the classical additive property of the variance function we obtain (93). ■

Now we can give two bounds for $e_{\varpi,m}(t_0)$. By using the Bienaymé-Chebyshev inequality, we obtain that for any real number $\gamma>0$:

$$Pr\left(|e_{\varpi,m}(t_0)-E[e_{\varpi,m}(t_0)]|\geq\gamma\sqrt{Var[e_{\varpi,m}(t_0)]}\right)\leq\frac{1}{\gamma^2}.$$

Then,

$$Pr\left(|e_{\varpi,m}(t_0)-E[e_{\varpi,m}(t_0)]|<\gamma\sqrt{Var[e_{\varpi,m}(t_0)]}\right)>1-\frac{1}{\gamma^2},$$

i.e.

$$e_{\varpi,m}(t_0)\in\,]M_l\,,M_h[\ \text{ with a probability }\ >1-\frac{1}{\gamma^2}, \tag{94}$$

where $M_l = E[e_{\varpi,m}(t_0)]-\gamma\sqrt{Var[e_{\varpi,m}(t_0)]}$ and $M_h = E[e_{\varpi,m}(t_0)]+\gamma\sqrt{Var[e_{\varpi,m}(t_0)]}$.

Remark 23 *Since the probability that the noise error contribution bound is not precise in general, it is better to calculate the distribution of the noise error contribution when it is possible.*

4.2 Error estimations of the first order derivative due to a gaussian noise

We assume in this section that the discrete noisy measurement is written as $y(t_i) = x(t_i) + C\varpi(t_i)$, where $C \in \mathbb{R}^+$, and the noise $\varpi(t_i)$ is a sequence of independent random variables with the same standard normal distribution ($\varpi(t_i) \sim \mathcal{N}(0,1)$). So the error $e_{\varpi,m}(t_0)$ is also a gaussian random variable. Since $e_{\varpi,m}(t_0) \sim \mathcal{N}(\hat{\alpha}, \hat{\beta})$ (with $\hat{\alpha} = E[e_{\varpi,m}(t_0)]$ and $\hat{\beta} = var[e_{\varpi,m}(t_0)]$), we have

$$e_{\varpi,m}(t_0) \overset{95.5\%}{\in} \left[\hat{\alpha} - 2\sqrt{\hat{\beta}}, \hat{\alpha} + 2\sqrt{\hat{\beta}}\right].$$

As the expected value of the noise is equal to zero, then according to Proposition 4.14, $E[e_{\varpi,m}(t_0)] = 0$. As the polynomial $p_{T,\Theta}(\cdot)$ is known, we can compute $var[e_{\varpi,m}(t_0)]$ by using Proposition 4.14 so that we can find out two bounds for the noise error contribution $e_{\varpi,m}(t_0)$.

In the next subsection, we are going to give the expression of some first order derivative estimators of x. We will also give the bounds for the associated noise error contributions.

4.2.1 Noise error contribution for some estimators

Let $N = n = 1$ and $\beta = \pm 1$ in Proposition 3.5. The minimal estimators are written in the following form:

$$\tilde{\dot{x}}_{t_0\pm}(k, \mu, \pm T) = \int_0^1 p_{k,\mu,\pm T}(\tau)\; y(\pm T\tau + t_0)\; d\tau, \tag{95}$$

where

$$p_{k,\mu,\pm T}(\tau) = \frac{1}{\pm T} \frac{(\mu + k + 3)!}{(k+1)!(\mu+1)!} \left((\mu + k + 2)\tau - (k+1)\right)(1-\tau)^{\mu}\tau^{k}. \tag{96}$$

The affine estimators $\tilde{\dot{x}}_{t_0\pm}(k, \mu, \pm T, 2, \xi)$ are given by (66):

$$\tilde{\dot{x}}_{t_0\pm}(k, \mu, \pm T, 2, \xi) = \lambda_1(\xi, k, \mu)\tilde{\dot{x}}_{t_0\pm}(k, \mu+1, \pm T) + \lambda_0(\xi, k, \mu)\tilde{\dot{x}}_{t_0\pm}(k+1, \mu, \pm T),$$

where $\lambda_1(\xi, k, \mu) = (k+3) - (\mu + k + 5)\xi$ and $\lambda_0(\xi, k, \mu) = 1 - \lambda_1(\xi, k, \mu)$.

According to (95) and (96), it reads

$$\tilde{\dot{x}}_{t_0\pm}(k, \mu, \pm T, 2, \xi) = \int_0^1 p_{k,\mu,\pm T,\xi}(\tau)\, y(\pm T\tau + t_0)\, d\tau, \tag{97}$$

where

$$p_{k,\mu,\pm T,\xi}(\tau) = \pm\frac{1}{T} \frac{(\mu + k + 4)!}{(k+2)!(\mu+2)!}(1-\tau)^{\mu}\tau^{k}(a_2\tau^2 + a_1\tau + a_0), \tag{98}$$

with $a_2 = (\mu + k + 3)(\mu + k + 5)\left[(\mu + k + 4)\xi - (k+2)\right]$, $a_0 = -(k+1)(k+2)\left[(k+3) - (\mu + k + 5)\xi\right]$, $a_1 = (k+2)\left[(2k^2 + 2k\mu + 12k + 5\mu + 16) - 2(\mu + k + 3)(\mu + k + 5)\xi\right]$.

If we replace y by ϖ in (95), then apply a numerical integration method, we obtain for the minimal estimators the noise error contributions denoted by $e^{\pm}_{\varpi,m}(t_0; k, \mu, \pm T)$. They are bounded by

$$|e^{\pm}_{\varpi,m}(t_0; k, \mu, \pm T)| \overset{95.5\%}{\leq} M^{\pm}_m(k, \mu, \pm T), \tag{99}$$

where $M^{\pm}_m(k, \mu, \pm T) = 2\sqrt{var[e^{\pm}_{\varpi,m}(t_0; k, \mu, \pm T)]}$. For the affine estimators, if we replace y by ϖ in (97), then by applying a numerical integration method, we have

$$|e^{\pm}_{\varpi,m}(t_0; k, \mu, \pm T, 2, \xi)| \overset{95.5\%}{\leq} M^{\pm}_m(k, \mu, \pm T, 2, \xi), \tag{100}$$

where $e^{\pm}_{\varpi,m}(t_0; k, \mu, \pm T, 2, \xi)$ are the noise error contributions, and $M^{\pm}_m(k, \mu, \pm T, 2, \xi) = 2\sqrt{var[e^{\pm}_{\varpi,m}(t_0; k, \mu, \pm T, 2, \xi)]}$.

Remark 24 *As for any $t_0 \in I$, $var[\varpi] = 1$, according to (93), we deduce that $var[e^{\pm}_{\varpi,m}(t_0; k, \mu, \pm T)]$ and $var[e^{\pm}_{\varpi,m}(t_0; k, \mu, \pm T, 2, \xi)]$ do not depend on t_0, so as to $M^{\pm}_m(k, \mu, \pm T)$ and $M^{\pm}_m(k, \mu, \pm T, 2, \xi)$. Moreover, we have*

$$M^+_m(k, \mu, T) = M^-_m(k, \mu, -T) \text{ and } M^+_m(k, \mu, T, 2, \xi) = M^-_m(k, \mu, -T, 2, \xi).$$

The central estimator can be given as follow by taking the formula in Proposition 3.11 with $N = n = 1$:

$$\begin{aligned}
\tilde{\dot{x}}_{t_0}(k, \mu, T) &= \frac{1}{2}\left(\tilde{\dot{x}}_{t_{0+}}(k-1, \mu, T) + \tilde{\dot{x}}_{t_{0-}}(k-1, \mu, -T)\right) \\
&= \begin{cases} \frac{1}{2}\int_0^1 p_{k-1,\mu,T}(\tau)\ \ (y(T\tau + t_0) - y(-T\tau + t_0))\ d\tau, & \text{if } k \geq 1, \\ \frac{1}{2}\int_0^1 p_{-1,\mu,T}(\tau)\ \ (y(T\tau + t_0) - y(-T\tau + t_0))\ d\tau, & \text{if } k = 0, \end{cases}
\end{aligned} \tag{101}$$

where $p_{k-1,\mu,T}(\tau)$ is defined by (96), and $p_{-1,\mu,T}(\tau) = \frac{(\mu+2)(\mu+1)}{T}(1-\tau)^{\mu}$ is calculated according to Proposition 3.11.

If we replace y by ϖ in (101), then apply a numerical integration method, we obtain for the central estimator the noise error contribution denoted by $e_{\varpi,m}(t_0; k, \mu, T)$. It is bounded by:

$$|e_{\varpi,m}(t_0; k, \mu, T)| \overset{95.5\%}{\leq} M_m(k, \mu, T), \tag{102}$$

where $M_m(k, \mu, T) = 2\sqrt{var[e_{\varpi,m}(t_0; k, \mu, T)]}$.

Finally, taking the affine central estimator defined in (86) with $N = 2$ and $n = 1$, we have for $k \geq 1$

$$\begin{aligned}
\tilde{\dot{x}}_{t_0}(k, \mu, T, 2, \xi) &= \frac{1}{2}\left(\tilde{x}^{(n)}_{t_{0+}}(k-1, \mu, T, 2, \xi) + \tilde{x}^{(n)}_{t_{0-}}(k-1, \mu, -T, 2, \xi)\right) \\
&= \frac{1}{2}\int_0^1 p_{k-1,\mu,T,\xi}(\tau)\ \ (y(T\tau + t_0) - y(-T\tau + t_0))\ d\tau,
\end{aligned}$$

where $p_{k-1,\mu,T,\xi}(\cdot)$ is defined by (98). Moreover, we have also for $k \geq 1$:

$$\begin{aligned}
\tilde{\dot{x}}_{t_0}(k, \mu, T, 2, \xi) &= \frac{1}{2}\left(\tilde{x}^{(n)}_{t_{0+}}(k-1, \mu, T, 2, \xi) + \tilde{x}^{(n)}_{t_{0-}}(k-1, \mu, -T, 2, \xi)\right) \\
&= \lambda_1(\xi, k-1, \mu)\frac{\tilde{\dot{x}}_{t_{0+}}(k-1, \mu+1, T) + \tilde{\dot{x}}_{t_{0-}}(k-1, \mu+1, -T)}{2} \\
&+ \lambda_0(\xi, k-1, \mu)\frac{\tilde{\dot{x}}_{t_{0+}}(k, \mu, T) + \tilde{\dot{x}}_{t_{0-}}(k, \mu, -T)}{2} \\
&= \lambda_1(\xi, k-1, \mu)\tilde{\dot{x}}_{t_0}(k, \mu+1, T) + (1 - \lambda_1(\xi, k-1, \mu))\tilde{\dot{x}}_{t_0}(k+1, \mu, T).
\end{aligned}$$

Now, we are going to define $\tilde{\dot{x}}_{t_0}(k, \mu, T, 2, \xi)$ for $k = 0$. Denote it by $\tilde{\dot{x}}_{t_0}(\mu, T, 2, \xi)$. It can be given in the following form:

$$\tilde{\dot{x}}_{t_0}(\mu, T, 2, \xi) = \lambda_1(\xi, -1, \mu)\tilde{\dot{x}}_{t_0}(0, \mu+1, T) + (1 - \lambda_1(\xi, -1, \mu))\tilde{\dot{x}}_{t_0}(1, \mu, T). \tag{103}$$

According to (101), we have

$$\begin{aligned}
\tilde{\dot{x}}_{t_0}(\mu, T, 2, \xi) &= \frac{1}{2}\int_0^1 \left(\lambda_1(\xi, -1, \mu)p_{-1,\mu+1,T(\tau)} + (1 - \lambda_1(\xi, -1, \mu))p_{0,\mu,T(\tau)}\right)(y(T\tau + t_0) - y(-T\tau + t_0))\ d\tau \\
&= \frac{1}{2}\int_0^1 p_{-1,\mu,T,\xi}(\tau)\,(y(T\tau + t_0) - y(-T\tau + t_0))\ d\tau,
\end{aligned}$$

where $p_{-1,\mu,T,\xi} = \frac{1}{T}(\mu+2)(\mu+3)\left(((\mu+3)\xi - 1)(\mu+4)\tau - 2(\mu+4)\xi + 3\right)(1-\tau)^{\mu}$.

If we replace y by ϖ in the affine central estimator, then apply a numerical integration method, we obtain the noise error contribution denoted by $e_{\varpi,m}(t_0; k, \mu, T, 2, \xi)$. It is bounded by:

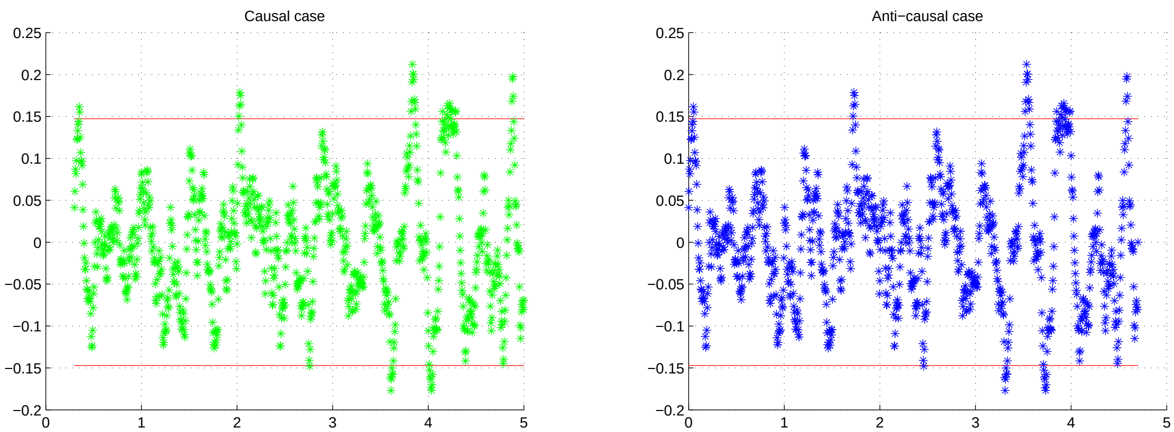
$$|e_{\varpi,m}(t_0; k, \mu, T, 2, \xi)| \overset{95.5\%}{\leq} M_m(k, \mu, T, 2, \xi), \tag{104}$$

where $M_m(k, \mu, T, 2, \xi) = 2\sqrt{var[e_{\varpi,m}(t_0; k, \mu, T, 2, \xi)]}$.

It can be seen that all the bounds values depend on the parameters k, μ, T or ξ as well as on the numerical integration method. We are going to analyze their influence on the noise error contribution.

4.2.2 Analysis of the influence of the parameters

Let $y(t_i) = x(t_i) + C\varpi(t_i)$ be the discrete noisy measurement observed in the time interval $]0, 5[$ with a sampling period $T_s = 1/200$ and with $C = 0.05$. For the minimal estimators, we compare in Figure 3(a) and Figure 3(b) the simulated noise error contributions $e^{\pm}_{\varpi,100}(t_0; 0, 0, \pm 0.3)$ and the calculated error bounds $M^{\pm}_{100}(0, 0, \pm 0.3)$ defined by (99). We use here the Trapezoidal rule with $m = 60$ ($T = 0.3$), $k = 0$ and $\mu = 0$. As we can see, these bounds are near optimal.



(a) $e^{-}_{\varpi,60}(t_0; 0, 0, -0.3)$ and $M^{-}_{60}(0, 0, -0.3) = 0.1472$

(b) $e^{+}_{\varpi,60}(t_0; 0, 0, 0.3)$ and $M^{+}_{60}(0, 0, 0.3) = 0.1472$

Figure 3: Comparisons between the noise error contributions and their bounds

From (99), after some calculations, we deduce that for any fixed value of m the local minimum of $M^{+}_{m}(k, \mu, T)$ is obtained for $k = \mu = 0$. For $k = \mu = 0$ and $m \in \{30, 40, \cdots, 170\}$, we calculate in Figure 4 the values of the error bound $M^{+}_{m}(0, 0, T)$ according to different numerical integration methods. As we can see, the Trapezoidal rule seems to be optimal. Consequently, we need to take into account this fact so as to optimize the parameters for minimizing the global noise error contribution.

For the affine estimators, according to [35], we should take $\lambda_0(\xi, k, \mu) = \lambda_1(\xi, k, \mu) = 0.5$ for minimizing $var[e^{\pm}_{\varpi}(t_0; k, \mu, \pm T, 2, \xi)]$ the variances of the noise error contributions. When k and μ are fixed, we should take ξ such that $\lambda_1(\xi, k, \mu) = (k + 3) - (\mu + k + 5)\xi = 0.5$, *i.e.*

$$\xi = \frac{k + 2.5}{k + \mu + 5}. \tag{105}$$

It is clear that ξ is increasing with respect to k, decreasing with respect to μ and is independent of T. According to (105) we obtain that $M^{\pm}_{m}(k, \mu, \pm T, 2, \xi)$ only depend on three parameters k, μ and m. After some calculations, we obtain in Table 1 the optimal parameters k, μ, ξ and T for some different numerical integration methods, with which the gaussian error effect can be minimized for these estimators.

Estimators	Bounds	k	μ	ξ	T	Numerical method
affine estimator $\tilde{\dot{x}}_{t_0\pm}(k, \mu, \pm T, 2, \xi)$	$M^{\pm}_{m}(k, \mu, \pm T, 2, \xi)$	1	1	0.5	$\nearrow$	Trapezoidal rule
central estimator $\tilde{\dot{x}}_{t_0}(k, \mu, T)$	$M_m(k, \mu, T)$	0	0	$-$	$\nearrow$	Right rectangle
affine central estimator $\tilde{\dot{x}}_{t_0}(k, \mu, T, 2, \xi)$	$M_m(k, \mu, T, 2, \xi)$	1	2	0.5	$\nearrow$	Right rectangle

Table 1: Optimal parameters and numerical integration methods so as to reduce the noise error contribution

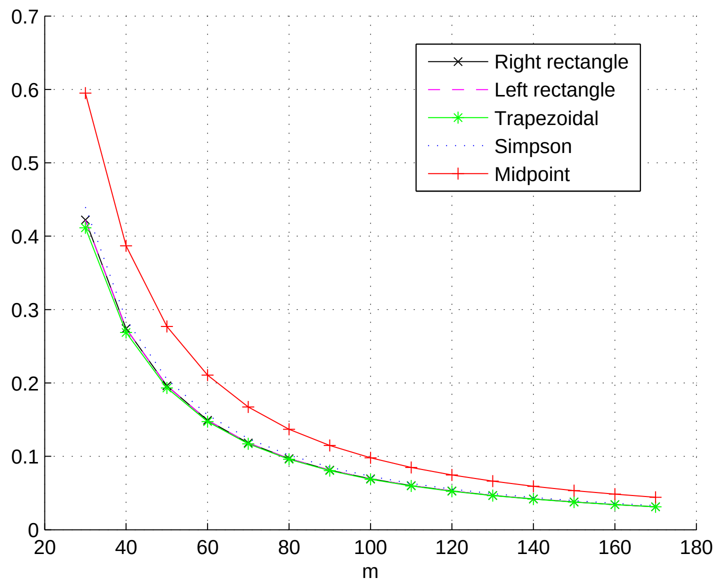


Figure 4: $M_m^+(0,0,T)$

Remark 25 *According to* (93) *in Proposition* 4.14 *and the expression of the bounds for the noise error contribution for each estimator, the optimal parameters and numerical integration methods found in this subsection for reducing each noise error contribution are independent to the variance of the noise.*

The global error for each estimator takes into account the bias term error, the noise error contribution and the error from numerical integration method (see Remark 22). In general, we need to choose the values of k and μ as small as possible so as to minimize the noise error contribution. This choice also allows us to neglect the numerical integration error. In the next subsection, we optimize the choice of k, μ, T and ξ for the affine causal estimator $\tilde{\tilde{x}}_{t_0-}(k,\mu,-T,2,\xi)$ from the study of the bounds of the total error.

4.2.3 Choice of parameters for the affine causal estimator

For the affine causal estimator $\tilde{\tilde{x}}_{t_0-}(k,\mu,-T,2,\xi)$, we investigate the total error which is the sum of the bias term error $e_{R_1}^-(t_0;k,\mu,-T,2,\xi)$ and the noise error contribution $e_{\varpi,m}^-(t_0;k,\mu,-T,2,\xi)$. The bounds for the bias term error $e_{R_1}^-(t_0;k,\mu,-T,2,\xi)$ are given in Proposition 3.10:

$$-\frac{(k+\mu+4)!}{(k+2)!(\mu+2)!}TM_l^- \leq e_{R_1}^-(t_0;k,\mu,-T,2,\xi) \leq -\frac{(k+\mu+4)!}{(k+2)!(\mu+2)!}TM_u^-,$$

where M_l^- is defined by (70) and M_u^- is defined by (71). The bounds for the noise error contribution $e_{\varpi,m}^-(t_0;k,\mu,-T,2,\xi)$ are given in (100):

$$-M_m^-(k,\mu,-T,2,\xi) \overset{95.5\%}{\leq} e_{\varpi,m}^-(t_0;k,\mu,-T,2,\xi) \overset{95.5\%}{\leq} M_m^-(k,\mu,-T,2,\xi).$$

Consequently, it yields

$$-\frac{(k+\mu+4)!}{(k+2)!(\mu+2)!}TM_l^- - M_m^-(k,\mu,-T,2,\xi) \overset{95.5\%}{\leq} e_{R_1}^-(t_0;k,\mu,-T,2,\xi) + e_{\varpi,m}^-(t_0;k,\mu,-T,2,\xi), \tag{106}$$

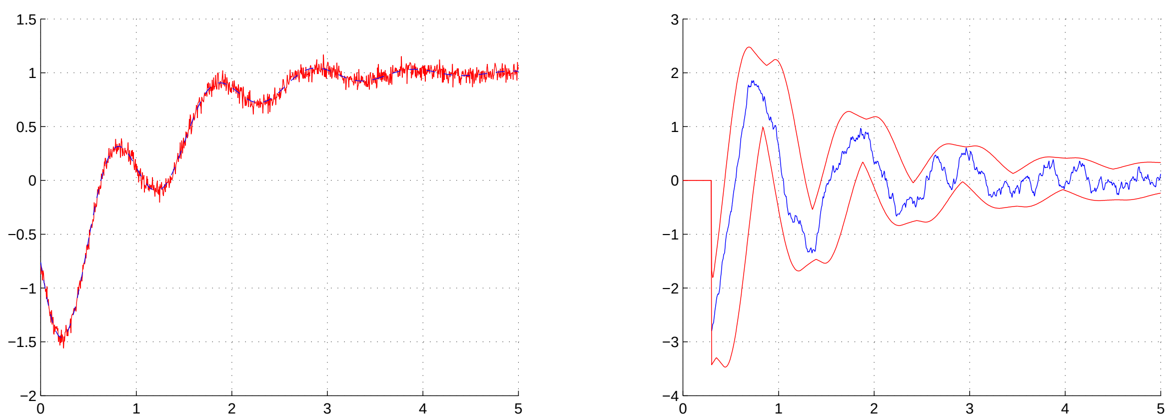
$$-\frac{(k+\mu+4)!}{(k+2)!(\mu+2)!}TM_u^- + M_m^-(k,\mu,-T,2,\xi) \overset{95.5\%}{\geq} e_{R_1}^-(t_0;k,\mu,-T,2,\xi) + e_{\varpi,m}^-(t_0;k,\mu,-T,2,\xi). \tag{107}$$

Now, let us consider the following noisy measurement

$$y(t_i) = \tanh(t_i - 1) + \exp(-t_i/1.2)\sin(6t_i + \pi) + C\varpi(t_i), \tag{108}$$

which is observed in the time interval $]0,5[$ with a sampling period $T_s = 1/200$, where the smooth signal (as in [35]) is $x(t_i) = \tanh(t_i - 1) + \exp(-t_i/1.2)\sin(6t_i + \pi)$ and the coefficient C is adjusted in such a way that the signal-to-noise ratio $SNR = 10\log_{10}\left(\frac{\sum|y(t_i)|^2}{\sum|C\varpi(t_i)|^2}\right)$ is equal to $SNR = 25$dB (see, e.g., [22] for this well known concept in signal processing). We can see this noisy measurement in Figure 5(a).

We explain here how to choose the parameters for the affine causal estimator $\tilde{\dot{x}}_{t_0-}(k,\mu,-T,2,\xi)$ in order to obtain a good estimation of the first order derivative of x. From the previous subsection we saw that the choice of $k=\mu=1$ and the use of the Trapezoidal rule help us to reduce the noise error contribution for $\tilde{\dot{x}}_{t_0-}(k,\mu,-T,2,\xi)$. With this choice, we use the affine causal estimator $\tilde{\dot{x}}_{t_0-}(1,1,-T,2,\xi)$ to estimate $\dot{x}$. We can see in Figure 5(b) the total error $\tilde{\dot{x}}_{t_0-}(1,1,-0.3,2,0.3)$ and its bounds in the case $T=\xi=0.3$.



(a) Noisy observation signal, $SNR = 25$dB

(b) Total error for $\tilde{\dot{x}}_{t_0-}(1,1,-0.3,2,0.3)$ and its bounds

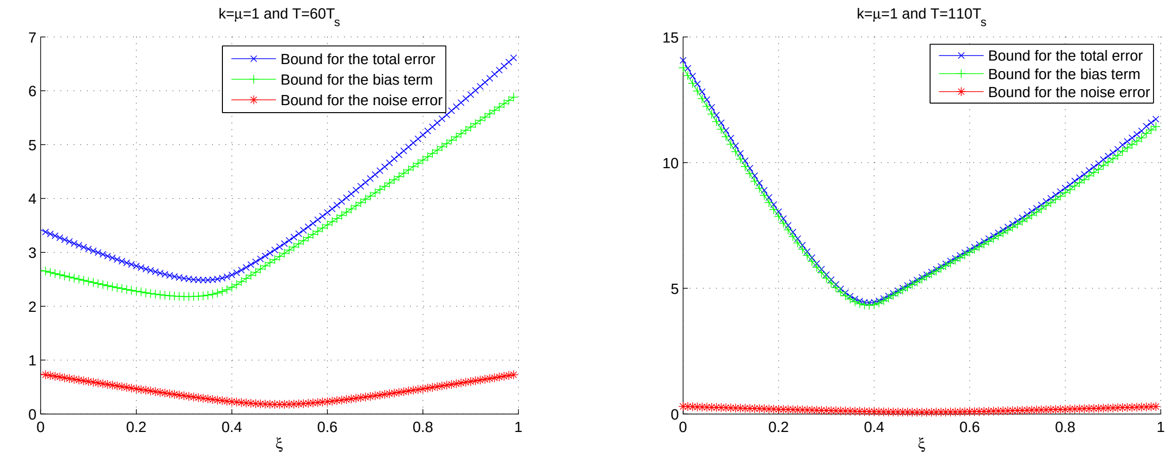
Figure 5: Noisy observation signal and the total error for $\tilde{\dot{x}}_{t_0-}(1,1,-0.3,2,0.3)$

We can observe that the total error reaches its maximum at $t_0 = 0.8$ and a local minimum at $t_0 = 1.3$. Figure 6 shows the values of the different upper bounds of the errors for $\tilde{\dot{x}}_{t_0-}(1,1,-T,2,\xi)$ at $t_0 = 0.8$ for any $\xi \in [0,1]$, where all the error bounds are positive. Figure 7 shows the values of the different lower bounds of the errors for $\tilde{\dot{x}}_{t_0-}(1,1,-T,2,\xi)$ at $t_0 = 1.3$ for any $\xi \in [0,1]$, where all the error bounds are negative.

It can be seen that he bounds for the noise error contributions are much smaller than the bounds for the bias term errors, especially when T increases. The extremum of the total error bounds are obtained when ξ is close to the smaller root of the second order Jacobi polynomial $P_2^{1,1}$ which is equal to 0.31.

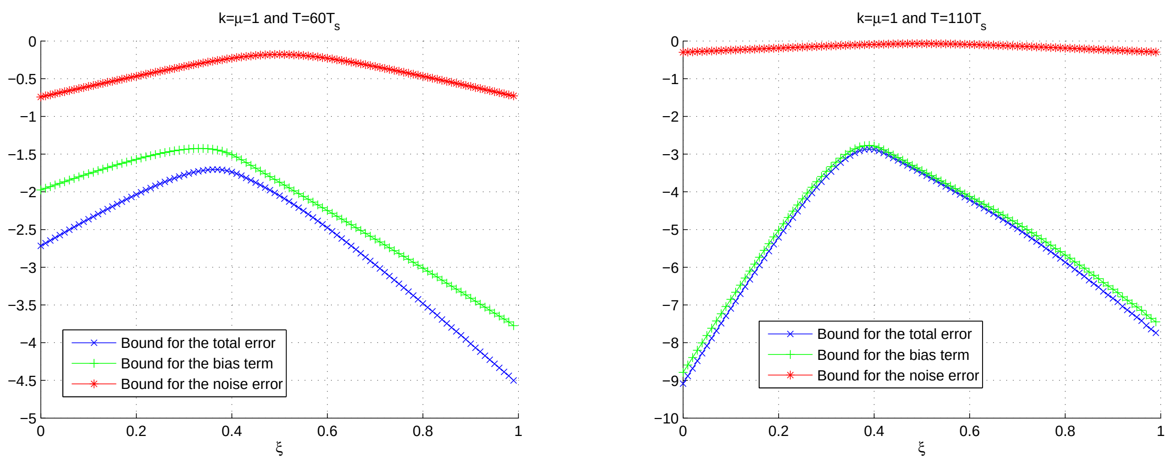
5 Comparison with some existing numerical differentiation methods

In this section, we compare some of the existing numerical differentiation methods with our methods. Let $\{y_j\}$ be a noisy measurement of an analytic signal x. We want to estimate its first order derivative. We assume that $\{y_j\}$ are



(a) In the case where $T = 60T_s = 0.3$ (b) In the case where $T = 110T_s = 0.55$

Figure 6: Upper error bounds for the errors for the estimation of $\dot{x}(t_0)$ at $t_0 = 0.8$



(a) In the case where $T = 60T_s = 0.3$ (b) In the case where $T = 110T_s = 0.55$

Figure 7: Lower error bounds for the errors for the estimation of $\dot{x}(t_0)$ at $t_0 = 1.3$

regularly spaced with a sampling period T_s (in practice such data are obtained using an AD converter or DSpace card which is working at a fixed sampling period).

Let us recall the general formula of the backward finite difference (BFD) and the central finite difference (CFD) differentiation schemes [6], [2] for $\tilde{\dot{x}}_j^{[p]}$, $p \in \{1, ..., 5\}$ which is the estimation of the first order derivative at point t_j

$$\tilde{\dot{x}}_j^{[p]} = \frac{1}{c_p n_m T_s} \sum_{i=-3}^{3} d_{i,p}\, y_{j+i\, n_m}, \;\; n_m \in \mathbb{N}^*, \tag{109}$$

where the coefficients are given in the following table

Scheme	p	$d_{-3,p}$	$d_{-2,p}$	$d_{-1,p}$	$d_{0,p}$	$d_{1,p}$	$d_{2,p}$	$d_{3,p}$	c_p
first order BFD	1	0	0	-1	1	0	0	0	1
second order BFD	2	0	1	-4	3	0	0	0	2
third order BFD	3	1	0	-9	8	0	0	0	6
first order CFD	4	0	0	-1	0	1	0	0	2
second order CFD	5	0	1	-8	0	8	0	-1	12

Remark 26 *The first order BFD (resp. first order CFD) scheme can be obtained by calculating the first order derivative of the Lagrange interpolating polynomial of degree* 1 *(resp.* 2*) at point* y_j*, which interpolates the noisy measurement at points* y_{j-n_m} *and* y_j *(resp.* y_{j-n_m}*,* y_j *and* y_{j+n_m}*). Thus, there are two sources of error for these schemes: the interpolation error and the noise error contribution. A simple formula for the interpolation error was shown in [30]. In [40] the author has described the required nodes used for interpolation in order to minimize the noise error contribution. However the noise was supposed to be bounded. In the next work, we will analyze the total error for the Lagrangian numerical differentiation with a random noise.*

The finite difference differentiation scheme is well-known as one of the most important instances of an ill-posed inverse problem [11] in the sense that small perturbations on the function to be differentiated may induce large changes in the derivatives. Generally, it has to be used in combination with a filtering device [5]. The moving average filter is the simplest digital filter to understand and use [28]. Now, let us recall the formula for the moving average method (see [28], [2] for details)

$$\bar{y}_j = \frac{1}{n_l + 1 + n_r} \sum_{i=-n_l}^{n_r} y_{j+i}, \;\; n_l, n_r \in \mathbb{N}. \tag{110}$$

The Savitzky-Golay filter can be considered as a generalized moving average [28]. It is based on the least squares polynomial fitting across a moving window within the data in the time domain. Let $n_L, n_R \in \mathbb{N}$ and $n_m \in \mathbb{N}^*$, then according to the Savitzky-Golay differentiation scheme [11], [12], for i fixed, $i \in \{-n_L, \cdots, n_R\}$, the estimation of the first order derivative at point $t_{j+i\,n_m}$ can be given by

$$\tilde{\dot{x}}_{j+i\,n_m} = \frac{1}{(n_m T_s)^n} p_n^T Y_{j,n_m}, \tag{111}$$

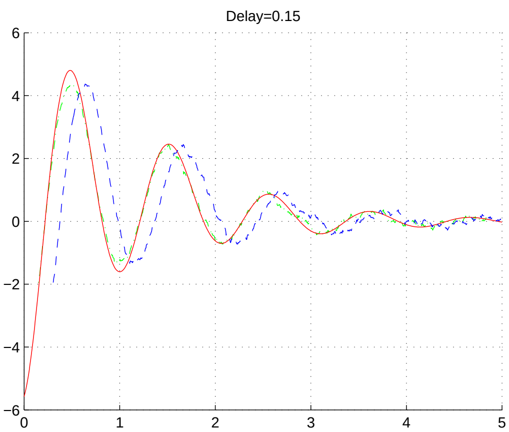
where $Y_{j,n_m} = \left(y_{j-n_m n_L}, y_{j-n_m(n_L-1)}, ..., y_{j+n_m n_R}\right)^T$, $p_n^T = \left(0, 1, 2i, \ldots, (N-1)i^{N-1}\right) B^\dagger$ and

$$B = \begin{pmatrix} 1 & -n_L & \cdots & (-n_L)^N \\ \vdots & \vdots & \vdots & \vdots \\ 1 & -1 & \cdots & (-1)^N \\ 1 & 0 & \cdots & 0 \\ 1 & 1 & \cdots & 1 \\ \vdots & \vdots & \vdots & \vdots \\ 1 & n_R & \cdots & n_R^N \end{pmatrix}.$$

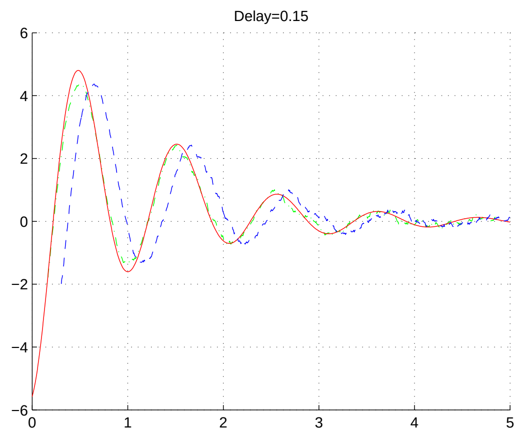
All the estimations of the first order derivative at point t_j obtained by the averaged finite difference differentiation schemes, the Savitzky-Golay differentiation scheme and the estimators defined in Section 3 can be considered as a dot product of a vector and the noise measurement in the moving window. This vector can be initialized as soon as the differentiation method and the associated parameters are chosen and it will be used in each moving window. Consequently, the problem of finding the best estimator is to find out the optimal vector.

Now let us consider the noisy measurement $y(t_j) = x(t_j) + C\varpi(t_j)$ defined by (108). We will compare the estimators introduced in Section 3 and the previous classical numerical differentiation methods when estimating $\dot{x}$ in the causal case and in the central case. To do so, we are going to calculate the total error variance $\int_0^5 e(\tau)^2 d\tau$ for each method.

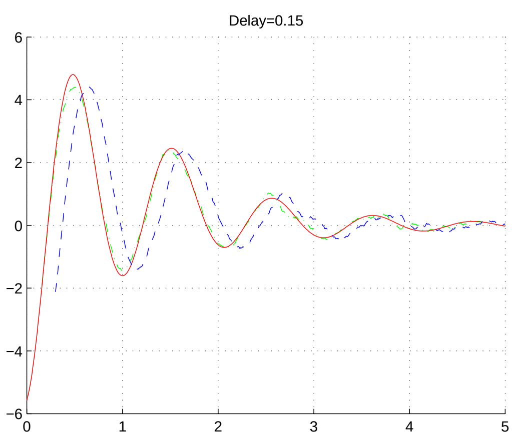
As the noise error contribution is discrete, we take $T_s \sum_{i=0}^{1000} e(\tau_i)^2$ as the approximation of $\int_0^5 e(\tau)^2 d\tau$. Moreover, we consider the SNR in each estimation. However, we have seen that it produces a delay in the minimal causal estimator $\tilde{x}_{t_0-}(k, \mu, -T)$ (see Remark 17) and in the affine causal estimator $\tilde{x}_{t_0-}(k, \mu, -T, 2, \xi)$ (see Remark 13). In order to obtain a better comparison in the causal case, the time-delayed estimations are shifted by the associated delay so that they become delay-free estimations, then we calculate the total error variance for the shifted estimations. We can see in Figure 8 how the time-delayed estimations are shifted. We can also see that the averaged BFD and the averaged Savitzky-Golay differentiation schemes are time-delayed with a delay which is close to half of the length of the moving window. In each figure, the solid line represents the exact derivative of x, the dashed line represents the time-delayed estimation and the dotted line represents the delay-free estimation (shifted estimation).



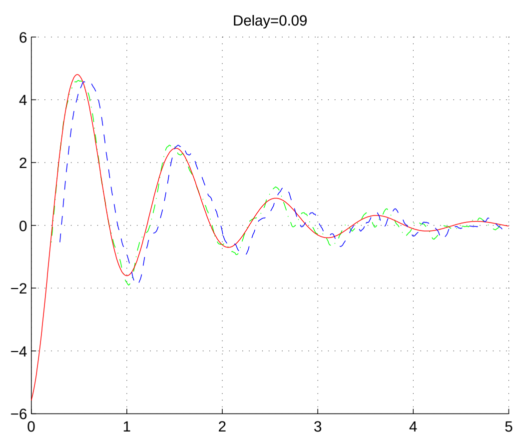
(a) Averaged BFD differentiation scheme



(b) Averaged Savitzky-Golay differentiation scheme



(c) Minimal causal estimator $\tilde{x}_{t_0-}(k, \mu, -T)$



(d) Affine causal estimator $\tilde{x}_{t_0-}(k, \mu, -T, 2, \xi)$

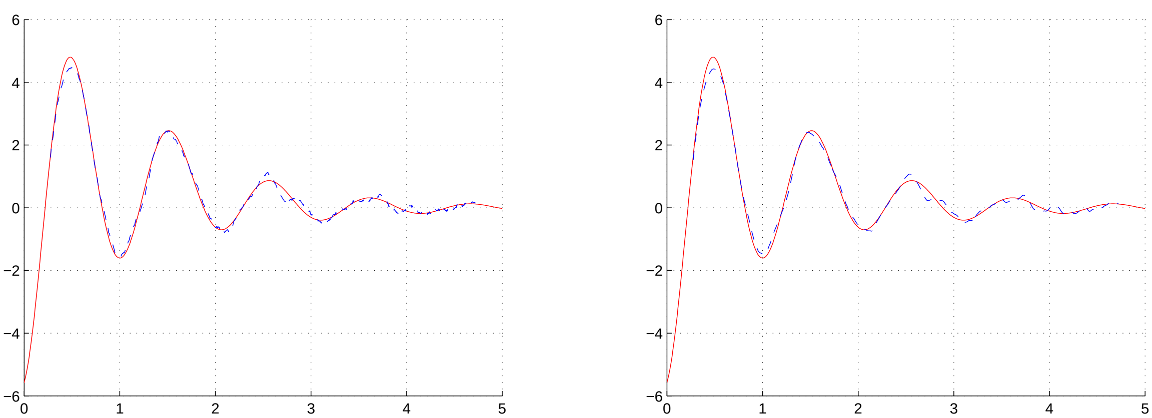
Figure 8: Time-delayed estimations and delay-free estimations with a moving window of length $T = 60T_s = 0.3$

We can see the comparison results in Table 2. In each grid, the first number represents the total error variance and the second number represents the SNR value for the associated estimation. All the values are calculated in the same time interval $[0.55 - \zeta, 5 - \zeta]$ with the same number of y_{t_j} for the moving estimative window, where ζ is the delay associated to each method. For the averaged BFD scheme, the parameters n_m and n_l are chosen such that the total error variance reaches its minimum. For the averaged Savitzky-Golay scheme, the parameters n_m, n_l and n_L are chosen similarly. For the causal estimator we choose $k = \mu = 0$ and $k = \mu = 1$ for the affine causal estimator where ξ is the smaller root of $P_2^{k,\mu}$ which is the second order Jacobi polynomial. We can observe that the minimal estimator and affine estimator can be considered as efficient filters. The minimal estimator is near optimal when $T = 60T_s$ and the affine estimator is near optimal with the smallest delay when $T = 110T_s$.

Value of T	first order BFD	Savitzky-Golay	Minimal Estimator	Affine Estimator
$T = 60T_s$	0.0965 | 54.8196	0.0792 | 52.1889	0.0774 | 53.9809	0.1583 | 38.7150
$T = 110T_s$	0.9130 | 67.4975	0.2660 | 67.9800	0.5027 | 73.2876	0.0530 | 57.5015

Table 2: Comparison in the causal case

Similarly, we can see in Figure 9 the estimations in the central case where these estimations are delay-free. In each figure, the solid line represents the exact derivative of x, the dashed line represents the delay-free estimation. In Table 3, we can see the comparison results in the central case. For the averaged CFD scheme, the parameters n_m and $n_l = n_r$ are chosen where the total error variance reaches its minimum. For the averaged Savitzky-Golay scheme, $n_L = n_R$. For the central estimator we choose $k = \mu = 0$ and $k = \mu = 1$ for the affine central estimator where ξ is the smaller root of $P_2^{k,\mu}$. We can observe that the central estimator is near optimal when $T = 60T_s$ and the affine central estimator is near optimal when $T = 110T_s$.



(a) Averaged CFD differentiation scheme

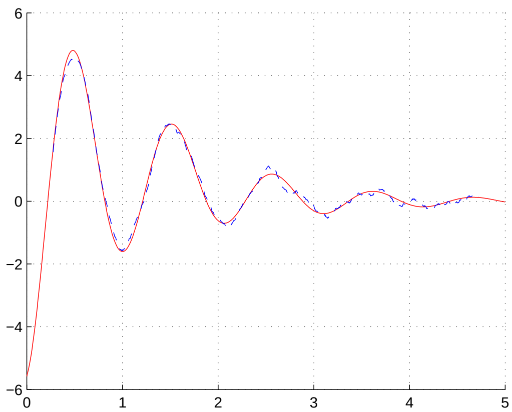
(b) Averaged Savitzky-Golay differentiation scheme

Value of T	second order CFD	Savitzky-Golay	Central Estimator	Affine Central Estimator
$T = 60T_s$	0.0557 | 49.0193	0.0628 | 51.3567	0.0471 | 52.6921	0.2892 | 36.4585
$T = 110T_s$	0.1388 | 67.4345	0.2121 | 67.2309	0.3024 | 71.1910	0.0619 | 50.9864

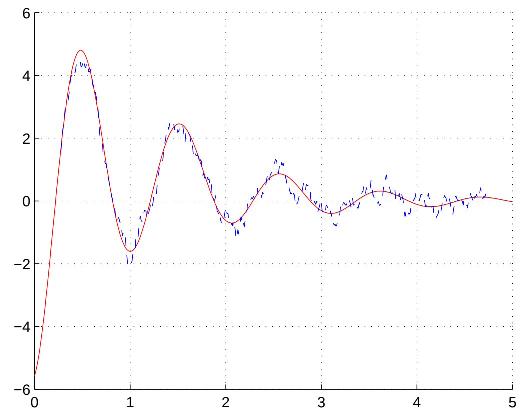
Table 3: Comparison in the central case

6 Conclusion

In the classical numerical differentiation methods, an interpolating polynomial or a least-squares polynomial is always used to approach a function, the derivatives of which we want to estimate. Then the derivatives of this approximation polynomial can be calculated as the derivative estimations which are in general the coefficients of the approximation



(c) Central estimator $\tilde{\tilde{x}}_{t_0}(k, \mu, T)$



(d) Affine central estimator $\tilde{\tilde{x}}_{t_0}(k, \mu, T, 2, \xi)$

Figure 9: Delay-free estimations with a moving window of length $T = 60T_s = 0.3$ in the central case

polynomial. More general interpolation and least squares approximations for numerical differentiation can be seen in [39] and [3]. In this paper, the recent algebraic parametric differentiation technic is presented: a N^{th} degree polynomial obtained from the Taylor series of an affine combination of a known function is given, then an annihilator rooted in [35] is applied to this polynomial in the operational domain, such that one of the coefficients of the polynomial is kept. The approximation of this remainder coefficient is taken as the estimation of the derivative. This derivative estimation contains two sources of error: the bias term error and the noise error contribution. The analysis for the noise error contribution of a known noise is given. Especially, the bias term errors are bounded for the minimal estimators and the affine estimators. We can give near optimal error bounds for these two estimators. These bounds are used so as to reduce the total error and give a guide to choose parameters. The error analysis for a numerical differentiation method is also given in [32] for unstructured and noisy data in the d-dimension case. Our central estimators can reduce bias term errors. We can use them to estimate the partial derivatives of a noisy data in two dimensions as in [32]. Finally, we show that these estimators are efficient by comparing with some of the improved classical numerical differentiation schemes.

Acknowledgment The authors would like to thank M. Mboub, C. Join and M. Fliess for interesting discussion on the topic.

References

[1] Al-Alaoui M.A.: A class of second-order integrators and low-pass differentiators. IEEE Trans. Circuits Syst. I 42(4), 220-223 (1995)

[2] Anderssen R. S., De Hoog F., and Hegland M.: A stable finite difference Ansatz for higher order differentiation of non-exact data, Bull. Austral. Math. Soc., 58, 223-232 (1998)

[3] Armentano M. G.: Error estimates in Sobolev spaces for moving least square approxiamtions, SIAM J. NUMER. ANAL. Vol. 39, No. 1, pp. 38-51 (2001)

[4] Barbot J. P., Fliess M., and Floquet T.: An algebraic framework for the design of nonlinear observers with unknown inputs. 46th IEEE Conference on Decision and Control (2007)

[5] Braci M. and Diop S.: On numerical differentiation algorithms for nonlinear estimation, in Proceedings of the 42nd IEEE Conference Dfl Decision and Control Maui, Hawaii USA, December 2003

[6] Brown RH, Schneider SC, and Mulligan MG: Analysis of algorithms for velocity estimation from discrete position versus time data, IEEE transactions on industrial electronics ISSN 0278-0046 CODEN ITIED6, vol. 39, no1, pp. 11-19 (12 ref.) (1992)

[7] Chen C.K. and Lee J.H.: Design of high-order digital differentiators using $L1$ error criteria. IEEE Trans. Circuits Syst. II 42(4), 287-291 (1995)

[8] Chitour Y.: Time-varying high-gain observers for numerical differentiation. IEEE Trans. Automat. Contr. 47, 1565-1569 (2002)

[9] Cullum J.: Numerical differentiation and regularization, SIAM J. NUMER. ANAL. Vol. 8, No. 2, June 1971

[10] Dabroom A.M. and Khalil H.K.: Discrete-time implementation of high-gain observers for numerical differentiation. International Journal of Control 72, 1523-1537 (1999)

[11] Diop S., Grizzle J. W., and Chaplais F.: On numerical differentiation algorithms for nonlinear estimation, in Pmceedings of the IEEE Conference on Decision and Conrml. New York IEEE Press Paoer CwOt876 (2000)

[12] Diop S., Grizzle J. W., Moraal P. E., and Stefanopoulou A.: Interpolation and Numerical Differentiation for Observer Design, University Michigan's College of Engineering Control Group Report Series, Report No. CGR-93-14, September 1993

[13] Duncan T.E., Mandl P., and Pasik-Duncan B.: Numerical differentiation and parameter estimation in higher-order linear stochastic systems. IEEE Trans. Automat. Contr. 41, 522-532 (1996)

[14] Fliess M.: Analyse non standard du bruit, C.R. Acad. Sci. Paris Ser. I, 342 797-802 (2006)

[15] Fliess M.: Critique du rapport signal à bruit en théorie de l'information, Manuscript (2007) (available at `http://hal.inria.fr/inria-00195987/en/`)

[16] Fliess M. and Diop S.: Nonlinear observability, identifiability and persistent trajectories. Proc. 36th IEEE Conf. Decision Control (1991)

[17] Fliess M., Mboup M., Mounier H., and Sira-Ramírez H.: Questioning some paradigms of signal processing via concrete examples, in Algebraic Methods in Flatness, Signal Processing and State Estimation, H. Sira-Ramírez, G. Silva-Navarro (Eds.), Editiorial Lagares, México, pp. 1–21 (2003) (available at `http://hal.inria.fr/inria-00001059/en/`)

[18] Fliess M. and Sira-Ramírez H.: Reconstructeurs d'état, C.R. Acad. Sci. Paris Ser. I, Vol. 338, pp. 91–96 (2004)

[19] Fliess M. and Sira-Ramírez H.: Control via state estimations of some nonlinear systems, Proc. Symp. Nonlinear Control Systems (NOLCOS 2004), Stuttgart, (2004) (available at http://hal.inria.fr/inria-00001096)

[20] Fliess M. and Sira-Ramírez H.: Closed-loop parametric identification for continuous-time linear systems via new algebraic techniques, *in* H. Garnier, L. Wang (Eds): *Identification of Continuous-time Models from Sampled Data*, pp. 363–391, Springer (2008) (available at `http://hal.inria.fr/inria-00114958/en/`)

[21] Fliess M. and Sira-Ramírez H.: An algebraic framework for linear identification, *ESAIM Control Optim. Calc. Variat.*, Vol. 9, pp. 151–168 (2003)

[22] Haykin S. and Van Veen B.: Signals and Systems, 2nd edn. John Wiley & Sons (2002)

[23] Herceg D. and Cvetković L.: On a numercial differentiation, SlAM J. NUMER. ANAL. Vol. 23, No. 3, June 1986

[24] Ibrir S.: Online exact differentiation and notion of asymptotic algebraic observers. IEEE Trans. Automat. Contr. 48, 2055-2060 (2003)

[25] Ibrir S.: Linear time-derivatives trackers. Automatica 40, 397-405 (2004)

[26] Ibrir S. and Diop S.: A numerical procedure for filtering and efficient high-order signal differentiation. Int. J. Appl. Math. Compt. Sci. 14, 201-208 (2004)

[27] Isaacson E. and Keller H.B.: Analysis of numerical methods, New York-Longon-Sydney: John Wiley and Sons, Inc (1966)

[28] Guiñón J.L., Ortega E., García-Antón J., and Pérez-Herranz V.: Moving Average and Savitzki-Golay Smoothing Filters Using Mathcad, September 3-7, 2007 International Conference on Engineering Education-ICEE (2007)

[29] Kochneff E.: Expansions in Jacobi Polynomials of Negative Order. Constructive Approximation vol. 13 n°4, 435-446, Springer (1997)

[30] Kranzer H.C.: An Error Formula for Numerical Differentiation, Numerische Mathematik 5, 439-442 (1963)

[31] Levant A.: Higher-order sliding modes, differentiation and output-feedback control. International Journal of Control 76, 924-941 (2003)

[32] Ling L.: Finding numerical derivatives for unstructured and noisy dada by multiscal kernel, SIAM J. NUMER. ANAL. Vol. 44, No. 4, pp. 1780-1800 (2006)

[33] Mboup M.: Parameter estimation for signals described by differential equations, Applicable Analysis, 88, 29-52, 2009.

[34] Mboup M., Join C., and Fliess M.: A revised look at numerical differentiation with an application to nonlinear feedback control. In: 15th Mediterranean conference on Control and automation (MED'07). Athenes, Greece (2007)

[35] Mboup, M., Join, C., and Fliess, M.: Numerical differentiation with annihilators in noisy environment, Numerical Algorithms 50, 4 439-467 (2009)

[36] Neves A., Mboup M., and Fliess M.: An Algebraic Receiver for Full Response CPM Demodulation, VI International telecommunications sysposium (ITS2006), Fortaleza-Ce, Brazil, (2006)

[37] Neves A., Miranda M.D., and Mboup M.: Algebraic parameter estimation of damped exponentials, Proc. 15^{th} Europ. Signal Processing Conf. - EUSIPCO 2007, Poznań (2007) (available at http://hal.inria.fr/inria-00179732/en/)

[38] Rader C.M. and Jackson L.B.: Approximating noncausal IIR digital filters having arbitrary poles, including new Hilbert transformer designs, via forward/backward block recursion. IEEE Trans. Circuits Syst. I 53(12), 2779-2787 (2006)

[39] Richter-dyn N.: Minimal interpolation and approximation in Hilbert spaces, SIAM J. NUMER. ANAL. Vol. 8, No. 3, September 1971

[40] Rivlin T. J.: Optimally stable lagrangian numerical differentiation, SIAM J. NUMER. ANAL. Vol. 12, No. 5, October 1975

[41] Roberts R.A. and Mullis C.T.: Digital signal processing. Addison-Wesley (1987)

[42] Su Y.X., Zheng C.H., Mueller P.C., and Duan B.Y.: A simple improved velocity estimation for low-speed regions based on position measurements only. IEEE Trans. Control Syst. Technology 14, 937-942 (2006)

[43] Szegö G.: Orthogonal polynomials, 3rd edn. AMS, Providence, RI (1967)

[44] Trapero J.R., Sira-Ramírez H., and Battle V.F.: An algebraic frequency estimator for a biased and noisy sinusoidal signal, Signal Processing, Vol. 87, pp. 1188–1201 (2007)

[45] Trapero J.R., Sira-Ramírez H., and Battle V.F.: A fast on-line frequency estimator of lightly damped vibrations in flexible structures, J. Sound Vibration, Vol. 307, pp. 365–378 (2007)

[46] Trapero J.R., Sira-Ramírez H., and Battle V.F.: On the algebraic identification of the frequencies, amplitudes and phases of two sinusoidal signals from their noisy sums, Int. J. Control, Vol. 81, 505-516 (2008)